

# Valediction from an old hand

**John Maddox steps down as editor of *Nature* from this issue, to be replaced by Dr Philip Campbell.**

L. J. F. (Jack) Brimble had been editor of *Nature* for roughly 22 years (for much of that time, in unwelcome harness with A. J. V. Gale) when he died almost exactly 30 years ago, alone in his apartment in London's Dolphin Square. I met him only once, over lunch one day. His manner was gruff, as if he had spent a lifetime in the army; I later learned that he was capable of great generosity to colleagues and even acquaintances. But he was evidently also given to bursts of irascibility: a few years earlier, when I had been working for the *Manchester Guardian*, he had removed me from the list of those to whom page-proofs of *Nature* were distributed in advance of publication. My transgression had been to write something that could be (and was) taken as critical of the journal. Then, without further intervention on my part, the tightly rolled brown-paper packages started arriving again.

Then Brimble died, and I had a visit from the late Maurice Macmillan MP. That was a minor embarrassment. I was then the Co-ordinator of the Nuffield Science Teaching Project, housed in a basement in one of the more dingy parts of Bloomsbury. My office was a partitioned box with no windows, and we had no decent coffee cups (nor coffee). But Macmillan had said that he had wanted to talk about *Nature*, which we did. I can imagine, but cannot remember, what I said: probably "Strange to see a weekly journal with so little in the way of news", and I would probably also have boasted of having found my first job (at the University of Manchester) from a classified advertisement in *Nature*.

But then the conversation became more serious, and Macmillan seemed to be asking whether I would be interested in Brimble's job. I remember a flush of ambivalence about the prospect (which is not the best way to win the confidence of potential employers). It had been more than a year since *Nature* had abandoned the practice of appending to research articles and notes the dates on which it had received them. That seemed worse than merely bad, which I explained to Macmillan. "How big is the backlog?", I asked. "I will arrange for it to be counted", he replied. When we next met, he produced the answer: 2,000-odd. (The number was precise, but I have forgotten all but the leading digit; I almost decided not to take the job, which was then firmly on offer.)

Poor Brimble was not entirely to blame. The sight of the office revealed his problem. It was an open-plan space without much of a plan. A window facing West ran 10

metres along the room and the broad window-ledge supported the famous backlog. That was arranged in piles, one for each month, providing a histogram of Brimble's problem, soon to be mine. There were fourteen monthly piles when I first saw them.

The staff was tiny, and evidently too small. There were two editorial people, one of whom had been recruited on Brimble's death; he had failed the first-year examinations at the medical school at which he had been enrolled. The more senior, and late Brimble's right-hand man, was Richard Fifield, who has been for many years the managing editor of *New Scientist*. The hard work at *Nature* in 1966 was done by two secretaries, Jill Baker (who later became a director of the Macmillan company owning *Nature*) and Mary Scallan (who is married to Nicholas Wade, the science editor of the *New York Times*). The roster was completed by two middle-aged men: Roy Lincoln, who handled the mail (of which there was a great deal) and D. W. Wilkinson (but everybody called him "Mr" Wilkinson). His job was to put illustrations in order for publication, which meant ordering type from a typesetter and sticking it alongside the ordinate and abscissa of any graph that might be published.

What I most vividly remember from that period is the sheer hard work, which seems inseparable from the job. As soon as I had shaken hands with the owner, then Mr Harold Macmillan (later the Earl of Stockton), I began to moonlight at the *Nature* office. We had a refereeing system of sorts in action within a few weeks, and began refusing manuscripts on the strength of negative opinions or their sheer length. Anxious to put something newsworthy in the journal, I began writing leading articles, much to the annoyance of the late Ronald Brightman, who lived in Cheshire and who worked from miscellaneous pamphlets sent to him in a parcel (by Roy Lincoln) every week. Soon we had our first newshound, a beginner journalist called Nigel Hawkes who is now the science editor of *The Times* (London). The last backlog manuscript was published only after eighteen months.

## Refereeing

The state of *Nature* in 1966 raises a raft of questions. Can, for example, a journal with such a reputation survive for long if scientific manuscripts are hardly ever refereed?

It is a good joke (which I have often used) that Watson and Crick's paper on the structure of DNA could not be



published now. It is only necessary to imagine what people would say if it reached them in the mail: "It's all model-building, just speculation, and such data as they have are not theirs but Rosalind Franklin's!" Some would complain that the sentence beginning, "It has not escaped our attention..." is entirely unsubstantiated, and must be an attempt to claim credit for developments in genetics that lie years ahead.

What happened when the paper reached the London office is easily imagined. Indeed, before sending it off, Sir Lawrence Bragg, the director of the Cavendish Laboratory in 1953, would have telephoned Brimble to warn him that the impending paper was important. Brimble might then have telephoned a few friends, perhaps including J. T. Randall, then the head of the Medical Research Council Biophysics Unit at King's College London, where both Franklin and M. H. Wilkins worked. It would not have taken them long to put together the account of Franklin's work, with its crucial demonstration of helical signatures in crystalline DNA. That paper appeared in the same issue as Watson and Crick, but after it.

These days, the idea that an editor should behave like that would be scorned. To have alerted the King's group to what was happening at Cambridge would be unethical. But whatever happened, rough justice was done; Wilkins shared the prize with Watson and Crick. (Rosalind Franklin's untimely death disembarassed the Nobel committees from having more than the statutory three candidates on their hands.) Who can complain?

### Confidentiality

This apocryphal reconstruction does not have to be true for it to illustrate the difficulties of the refereeing system now. One is the problem of the implied confidentiality. Journals ask referees to keep the contents of manuscripts confidential but have no way of policing their request. Referees are mostly honourable, but suppose one has been sent a paper whose conclusion is plausible and even arresting but unsubstantiated, and which the referee has recommended should not be published? Is it not then easier to talk about the idea or even to use it as the starting-point for an investigation of one's own? My guess is that the confidentiality of the refereeing process breaks down only when it is important that it should not.

For different reasons, during my first spell as editor of *Nature* (1966–73), there were two authors whose papers I decided never to send to referees. One was Louis Leakey, who repeatedly provoked otherwise reasonable colleagues to declare his paper not to be publishable and that he should spend the next three years working systematically through the fossils already stored in his many tea-chests; we would have missed much excellent palaeoanthropology if we had listened to them. The other was Fred Hoyle, who during the 1950s and 1960s entertained and instructed the readers of *Nature* on matters as different as the function of Stonehenge and the nature of the Universe. Most referees found him "unsound", often without saying why. But it does not require a referee to tell whether it is a reasonable

hypothesis that terrestrial life stems from interstellar bacteria. There remains a place for the occasional unrefereed article.

The other threat to the health and well-being of science is the fierce competition among scientists for money, mostly in the form of research grants and appointments. There is no doubt that the competitiveness of the US system, now being widely imitated elsewhere, has stimulated the US research enterprise to be more productive of good science than would have been expected. But it has also sacrificed large numbers of postdoctoral fellows on the altar of achievement. It seems a waste of young talent that the academic life should be so intellectually bruising. The obvious danger is that the yardsticks of attainment (impact factors, for example) used in the United States are unthinkingly transferred elsewhere. But that seems all too likely.

Civility has been the other casualty of competitiveness, especially in the United States. Endless disputes about individual priority for discoveries are one almost trivial index of the underlying trouble. One incident in the past few months sticks in the mind. An author found that an article submitted to *Nature* was returned on the grounds that it answered only half the question posed, but was then invited to resubmit it when another article dealing with the other half of the question arrived by coincidence. The original author professed himself delighted, but failed to keep a promise to return his manuscript quickly. After a month's delay, it appeared that it was already in the press with another journal. The original author did not have to share the limelight with another.

The English language is another casualty. It used to seem that *Nature's* contributors wrote clearly, but no longer. But has not science itself become much more intricate? That is true, but not so much as to justify the obscurity of most scientific texts. Nor is it that authors have become lazy; the accompanying figures are always excellently prepared, as are the typescript pages. The obscurity of the literature now is so marked that one can only believe it to be deliberate. Do people hide their meaning from insecurity, for fear of being found out or, in the belief that what they have to say is important, to hide the meaning from other people?

### Misconduct

The scientific literature also suffers in ways other than the language it contains. Misconduct of some overt kind is on the rise. We all know why that is. Reputations rest on publications as never before, as do promotions and research grants. Evidently the pressure intensified remarkably between 1973 and 1980. During the spell 1966–73, there was hardly a case of scientific misconduct to worry about. By evil accident, one of the first pieces of mail I read when I returned in 1980 was a warning, from a British academic against the activities of a Jordanian called Alsabti, operating at Houston, Texas, at the time. It turned out that he had been publishing about one paper a month in undistinguished journals, citing very distinguished people as his co-authors. As the world knows, it is only now that the torrent

of scandal is slowing down.

There can be no more important goal for the research community in the next few years than to cut the link between publications and success that this pattern of misdeemeanour implies. The extreme cases of fraud are less alarming than the present inversion of the link between the publications records of honest scientists and their success. If we agree that successful people publish a great deal because they have a lot to say, does it follow that people who publish a great deal are successful? Of course not, especially because journals have no means of ensuring that each published gobblet is of similar quality.

I believe that it is still true that almost every issue of every scientific journal contains something that will be surprising, even illuminating. But it is hard to avoid the conclusion that most of what is published need not be published urgently, or where it appears or even at all. That does not mean that the research enterprise is bankrupt, but that people cannot on the one hand complain that they cannot keep up with the literature and, on the other, insist that their survival depends on publication.

## Discovery

None of this implies that the past 30 years have been one long battle with referees and aggressive authors. On the contrary, they have been filled with huge interest. Within a few months of *Nature's* centenary in 1969, for example, we had Mark Ptashne's discovery of the *lac* operon (the first genetic regulatory element to be recognized), and the independent discovery by David Baltimore and Howard Temin that the enzyme reverse transcriptase can convert RNA into genetically equivalent DNA. And there was the discovery of the first pulsating star (now one of thousands). By 1972, people were talking of genetic engineering (or at least of a moratorium in the field for a year or so).

To compare the speed with which understanding is being deepened in the life sciences with what happened in physics in the 1920s is probably flattering to physics. Can there ever have been a time when there have been so many people pushing at an open door as there are now in the parts of biology derived from what was once called molecular biology? Sometimes it seems as if the intelligent statement of a question evokes the answer without much further effort.

With genome technology all the rage, it is natural that there should be a widespread expectation of still further improvements of health-care, at least in the rich countries of the world. That may be over-optimistic. The only certainty is that when the functions of newly discovered genes are better understood, there will be great opportunities for the pharmaceutical companies. Manipulating single genes by gene therapy may yet prove feasible, but manipulating groups of genes that are widely separated but must act in concert will be a headache for years to come. But these are not the important prizes. One is the improvement of crop plants by genetic technology and the other is the understanding of human (and mammalian) evolution that will flow from comparative analysis of the genomes of related species. It would be good to see the Human Genome

Diversity Project given a fair wind in the years ahead.

What will happen to the physical sciences, now almost dormant in comparison with biology? Nobody should be disheartened. The overdue upheaval in cosmology will be immensely enlivening for us all, so too will be the understanding that must soon come of why the supposedly fundamental particles of matter in the Universe have just the properties observed, and not some others. These elements of the future are more easily described than they will be brought about, but some puzzles are not even simply stated. The old conundrum: how does the brain work? is not now as intractable as a few years ago, but the question of when life began on the surface of the Earth (or elsewhere) is as unapproachable as ever, enthusiasm for the problem notwithstanding.

Meanwhile, there is technology. The nanochip is on the way, as are the self-assembling electronic components modelled on what happens in cells, brain cells in particular. In due course there will be ways of assembling single large molecules that meet every specification. What use this will be to the world at large is clear only to the extent that the past has shown already what benefits flow from developments of just that kind. *Nature* will no doubt have a big part to play in these developments, as it is already doing.

*Nature's* other role, as a critic of the research community's manners and of the frequent misuse of science and technology, is something else again. The fuss that has been stirred up by recent developments in genetics will not quickly melt away. The crying need is for a better public understanding of what is being done, and why, together with a set of international agreements on the proper use of these new technologies. Global warming is here to stay, at least as a problem and probably as a phenomenon. Educating the young in novel ways, always with us, will become more taxing. And there remains the plight of the poor in rich countries and the poor in countries where all but a few are like themselves.

The scientific community needs several means of representing its legitimate interests to the governments and other public bodies, most of which are indifferent to them. The British government's lackadaisical transfer of the newish (for a government department) Office of Science and Technology to the Department of Trade and Industry is a particularly scandalous way in which to treat the scientific enterprise, especially when everybody in government is crying that the function of science is to create wealth, and quickly please. But there may be worse to come. Most industrialized countries are busily cutting spending so as to balance their books; science and higher education are easy targets, at least while they do not answer back. *Nature* has traditionally played a part in battles of this kind, and will no doubt continue in that vein.

For my part, I wish *Nature* and my successor, Philip Campbell, the success that they deserve, and thank my colleagues (past and present) for good fellowship over many years and a host of readers and contributors for correspondence that has always been illuminating and often entertaining.

John Maddox



# Climate panel confirms human role in warming, fights off oil states

**London.** The scientific consensus on global warming took a further step forward last week when the United Nations' main advisory panel endorsed a scientific report concluding that "the balance of evidence suggests that there is a discernible human influence on global climate".

But even this conclusion, reached by Working Group 1 of the Intergovernmental Panel on Climate Change (IPCC), remains controversial, this time over an instruction to authors of the report to make last-minute revisions in the text, a move that some claim might contravene the IPCC's guidelines.

The call for modifications stemmed from bitter argument during a three-day meeting in Madrid as experts from IPCC countries sought to agree a summary of the working group's report collating scientific data on climate change. Delegates from two oil-producing nations, Saudi Arabia and Kuwait, both of which feel threatened by any curbs on the use of oil, argued at length that the summary should contain language emphasizing what they claimed to be the 'uncertain' nature of recent evidence pointing to a human effect on climate.

Other delegates disagreed. But after eight and a half hours of deadlock, they conceded to a footnote in the summary stating that "two countries" considered that recently published evidence showing a correspondence between the results of climate models and observations of climate change should be regarded only as "preliminary".

The footnote request was granted partly on the grounds that one chapter of the full report also suggests that some of the new studies on climate change have large caveats and have yielded preliminary evidence of a human effect on climate.

Mohammad Al-Sabban, an adviser to the Saudi minister of petroleum based in Jeddah and head of the kingdom's delega-

tion to the IPCC, says he agrees that the evidence for global warming is beginning to accumulate. "But to attribute this solely to human activity is too simplistic."

But Tom Wigley, senior scientist at the US National Center for Atmospheric Research in Boulder, Colorado, a lead author of the chapter concerned, says the word 'preliminary' is used in the chapter in

IMAGE  
UNAMBIGUOUS  
FOR A COVERT  
REASON  
REASONS

**Taking cover: few dispute that rainfall will increase.**

a different context. This word, he says, implies that evidence for a human effect on climate change is initial, but clear and unambiguous. "It does not mean that evidence of human influence on global climate is uncertain." Wigley is now making clarifications to the report to remove any ambiguity. "We did not realize how this word could be misinterpreted," he says.

It remains unclear whether modifications are permitted under IPCC guidelines at such a late stage. According to the guidelines, the contents of each chapter need to be peer-reviewed before publication. The government and scientific review process ended earlier this year. And the report will now be considered for adoption as part of the IPCC's Second Assessment Report at a meeting of scientists and government representatives in Rome next week.

Sir John Houghton, co-chair of IPCC Working Group 1, denies that IPCC rules are being broken. "The rules state clearly that background material must be consistent with the policy-makers' summary."

Houghton maintains that the scientific integrity of the IPCC has not been dented by the re-write. "This is not a fudge or a cover-up," says Houghton. "The compromise has been made on presentational points. If the science had been changed, I would have abandoned the process."

The summary itself has been widely praised as an authoritative document that should help to provide a firm basis for negotiations to reduce emissions of greenhouse gases at the next UN conference of the parties to the Framework

Convention on Climate Change.

The summary points out that "recent years have been among the warmest since 1860". It adds that global mean temperatures have increased by between 0.3 degrees and 0.6 degrees Celsius since the late nineteenth century, and that global sea levels have risen by between 10 and 25 cm over the past century.

Future years, it says, will be characterized by weather extremes of cold and drought. Temperatures are expected to rise by a further one degree Celsius, and sea levels by another 15 cm by 2100. If carbon dioxide emissions are maintained at their 1994 levels, according to the report, atmospheric concentrations of the gas would reach about 500 parts per million by volume, nearly twice their pre-industrial concentrations.

Despite the strength of these conclusions, some non-governmental organizations are concerned about suggestions that the Saudi and Kuwaiti delegates, backed, it is claimed, by the Climate Council, a major US fossil-fuel lobbying group, tried to upset the discussions and are partly responsible for large sections of the report being removed from the final summary through lack of negotiating time at last week's meeting.

Instead of the original 40-page draft document, delegates were able to approve only an 11-page version that was initially intended as an executive summary. The full summary, which contains important graphs, tables and numerical references, will now be added as a technical annex.

Both the Saudi and Kuwaiti delegates were also criticized for failing to attend key contact group meetings — and then holding up plenary sessions by tabling extended amendments. "The situation is too serious for countries such as Saudi Arabia and Kuwait to continue trying to subvert the IPCC process," says Merylyn McKenzie Hedger, climate policy officer for the Worldwide Fund for Nature.

But Al-Sabban rejects allegations that his group tried to obstruct the negotiating process. The summary, he says, "is now balanced and cautious enough not to mislead policy-makers". Nevertheless, he says, Saudi Arabia is entitled to protect its own interests. "Saudi Arabia's oil income amounts to 96 per cent of our total exports," he says. "Until there is clearer evidence of human involvement in climate change, we will not agree to what amounts to a tax on oil."

**Ehsan Masood**

## Argentina may host cosmic ray detector

**Paris.** Argentina has been chosen as the proposed site of the southern half of a giant, twin-location cosmic-ray observatory being put forward by an international panel of physicists. Meeting in Paris two weeks ago at the United Nations Educational, Scientific and Cultural Organization (Unesco), which has already provided \$100,000 for a design study of the planned observatory, scientists from 19 countries decided that the detector, which would cover a total area of 3,000 square kilometres, should be built in Argentina's Mendoza province. □

# Scientists 'must assert their independence'

**London.** Sir Michael Atiyah, one of Britain's leading mathematicians, ended his four-year term of office as the president of the Royal Society last week with an outspoken call for scientists "to criticize the establishment where necessary", and to "demonstrate that independence of thought really is the hallmark of a scientist".

True to his own counsel, he then attacked the support given since the Second World War by successive British governments to an independent nuclear deterrent. Atiyah claimed that this has been the root not only of the country's economic decline over that period but also of the hostile attitude of the general public towards science and the scientific community. He also expressed concern about the continuing role of scientists in the international arms trade.

At one point, Atiyah implicitly compared the dangers of the government's lack of support for basic science to those of Stalin's attitude towards the arts. But, in a move apparently intended to indicate political even-handedness, he also criticized hints in a recent Labour Party document that the society's receipt of public funds could be used as a possible lever to pressure it into appointing more women as fellows.

Atiyah was making his last formal statement before handing over to his successor, Sir Aaron Klug, the director of the Medical Research Council's Laboratory for Molecular Biology in Cambridge. He said that his



**Atiyah: critical of the UK's nuclear deterrent.**

remarks were prompted by concern that the scientific community is "in danger of losing our way and our identity" in the "semi-political" world in which it has to operate.

Referring, for example, to this year's fiftieth anniversary of the dropping of the atomic bomb, he emphasized how, for many scientists, this had demonstrated that "the ivory tower was no longer a sanctuary" and the public felt "atomic bombs were a menace and the scientists were responsible".

The crucial question now faced by scientists is how to conduct relations with government and industry so as to regain the public's confidence. Here some humility is needed. "It is no use complaining that the public is simply ill-informed and needs re-educating," said Atiyah. "We have to examine our own position and see whether any of the criticisms levelled against us are valid."

Even though scientists may constantly

exercise a "benign influence" in the corridors of power, this will not impress a sceptical public. "Scientists are too often thought of as a secretive élite, a sinister part of the establishment, part of 'them' not part of 'us'," said Atiyah. "The only way to break down this suspicion and distrust is for scientists to speak out openly and freely."

For example, Atiyah predicted that, on the basis of comparisons with the recent histories of Germany and Japan, "history will show that the insistence on a UK nuclear capability was fundamentally misguided, and a significant factor in our relative economic decline over the past 50 years".

At the same time, despite promises of a 'peace dividend' at the end of the Cold War that was supposed to see the shift of resources from military to civilian research, Atiyah said that he had "failed to detect any conscious policy on the redistribution of scientific resources".

Meanwhile, the export of conventional armaments raises continued ethical problems for scientists. Atiyah said that he acknowledges the economic and employment advantages from such sales. "But as a scientist, I cannot by my silence condone a policy which uses the scientific skills of this country to export potential death and destruction to poorer parts of the world, where their scarce resources would be better employed on food and health."

Turning to contemporary science policy issues, Atiyah also warned that turning off resources for so-called 'blue-skies' research will encourage the creative scientific intellects needed for the future to migrate to other lands or other occupations.

The dangers of political pressure, he suggested, also emerge in proposals that the government should take steps to require the society to increase its women fellows. Atiyah pointed out that a number of steps have recently been taken to assist women scientists at various stages of their careers — including a new fellowship scheme named after the Nobel prizewinner Dorothy Hodgkin — but said that it will take time for these measures to produce a larger flow of women candidates.

**David Dickson**

● The holder of Britain's first chair in the public understanding of science claimed last week that there is a need to shift from a "top down" approach, based on the premise that the main need is merely to propagate scientific information more widely, to one that is "bottom up" as based on a dialogue between scientists and the public.

Delivering his inaugural address at Imperial College, London, John Durant — who is also an assistant director of London's Science Museum — said that those engaged in developing the public understanding of science need to adopt a new agenda, namely the creation of trust between the two sides.

**Sally Lehrman**

## CIA's psychic spies under scrutiny

**San Francisco.** A \$20-million 20-year US government programme to employ and evaluate paranormal spying techniques has yielded results which, while not providing any firm evidence of psychic phenomena, can still be described as equivocal, according to two analysts who have reviewed the previously classified work.

Their conclusions are based on recently released data on the way various US government agencies secretly used up to six psychics at a time to help locate hostages, track down alleged terrorists and help drug enforcement officials. Experiments included studies on precognition, clairvoyance and "remote viewing", in which psychics visualized distant or hidden locations or objects.

The reviewers agreed that the results were statistically significant. But they disagreed on how they should be interpreted. Jessica Utts, a statistics professor at the University of California at Davis who is an advocate of parapsychology research, said they showed psychic phenomena as having had a small to medium-sized effect. But her collaborator, Ray Hyman, a psychology professor at the University of Oregon in Eugene, said that "inexplicable statistical

departures from chance" were a far cry from compelling proof.

In the remote viewing studies, for example, a 'viewer' accurately drew a California windmill farm and the footbridge at a wildlife refuge visited by a 'sender' partner as many as 50 miles away.

The Central Intelligence Agency had asked Utts and Hyman to evaluate the project for the Senate Appropriations Committee. The programme, which employed three psychic agents as recently as July, has raised questions about appropriate government spending on research. The studies began as a response to concerns that a 'psychic gap' in military and intelligence prowess had developed with the Soviet Union.

Utts and Hyman agreed that the first era of research was problematic, with no controlled experiments and selectively chosen research results. But, they said, protocols improved greatly by the late 1980s. Utts, who had participated in some of the studies, said that future research should concentrate on how a psychic sense may work. But Hyman said more work is needed to determine whether psychic phenomena exist.

## Report on US nuclear repository 'needs effective peer review'

Washington. An independent panel of scientists has harshly criticized a key part of the process which the US Department of Energy (DoE) is using to decide about the use of Yucca Mountain in Nevada as a repository for the long-term storage of nuclear waste.

The department is preparing a series of six Technical Basis Reports (TBRs). These are supposed to present the science that will underpin its final decision, due in 1998, on whether Yucca Mountain is fit for the purpose.

But in a stinging rebuttal of the first such report — on surface characteristics, hydrology and erosion — a panel of the National Research Council (NRC) says that the DoE failed to include the right data, and did not present it in the right way.

"The DoE report is deficient," says Ernest Smerdon, dean of engineering and mines at the University of Arizona, Tucson, and chair of the NRC panel. "It is not written clearly for a broad audience and does not include adequate supporting data for some of its analyses."

Jean Bahr of the geology and geophysics department at the University of Wisconsin at Madison, vice-chair of the panel, says that the criticism is of the report, not the underlying science.

"They don't have the support here that they'll need to convince people" that their decision on Yucca Mountain is the right one, she says. The NRC found that "many of the problems noted in this report could have been discovered and corrected had an effective peer review mechanism been in place".

The DoE reports were supposed to provide intellectual ballast for a decision on Yucca Mountain that is bound to be extremely controversial. The NRC's criticism may improve the prospects of their doing so in the long term — but may also fuel fears that DoE is mismanaging the entire project.

The report was jumped on by officials in the state of Nevada, which opposes the repository. Carl Johnson of the Nevada Nuclear Waste Project Office says that the findings bear out the state's view that the DoE has performed "a very narrow analysis, only using the DoE's own data".

The latest report only adds to the Yucca project's woes. The Department of Energy is now re-planning the project after the imposition of sharp cuts in its budget appropriation. Eight hundred contractor staff have been fired, the remaining five TBRs may be rolled into one, and the 1998 decision date may be moved back.

Colin Macilwain

## Planetary defence mission to yield pay-off for research

Washington. The US Air Force plans to launch a spacecraft in 1998 to approach several near-Earth asteroids and send microsatellites crashing into their surfaces to study the resulting impacts. But the National Aeronautics and Space Administration (NASA) may have mixed feelings about the military encroaching on its turf.

The mission, planned as a follow-on to the Clementine project that mapped the Moon last year but failed before reaching an asteroid (see *Nature* 367, 207; 1994), will test advanced satellite technology and provide data for a new Air Force programme designed for defence against the threat of Earth-crossing asteroids.

The 1996 defence appropriations bill approved by President Bill Clinton last week includes \$20 million for the Clementine 2 project. Its total cost is expected to be below \$80 million, according to Stewart, the head of the project, who was deputy manager for Clementine 1.

Nozette and the team from the Naval Research Laboratory (NRL) that built and flew the original Clementine have been searching for another assignment ever since a software glitch abruptly ended their first mission in May 1994. Proposals to NASA to fly a replacement for the lost Mars Observer and to send a spacecraft to the Moon as part of the space agency's Discovery series of low-budget planetary missions were both turned down.

Now the Air Force has agreed to fund Clementine 2 as part of a new "planetary defence" programme which, according to Colonel Simon Worden of the Air Force Space Command, has become "part of our mission statement".

The plan will be developed in detail over the next three months. It calls for a main spacecraft to carry several 10-kg microsatellites that will be sent hurtling into their targets at speeds of 5 to 10 km per second. Instruments on the main spacecraft will then observe the impacts and resulting craters.

The spacecraft could meet up with between two and six asteroids over the course of a two-year mission, says Nozette. Apart from the Clementine team at NRL, the project involves participation from the Air Force Phillips Laboratory and Lawrence Livermore National Laboratory.

Clementine 2 is seen as the first in a series of Air Force spacecraft that would eventually sample 20 to 30 Earth-crossing asteroids. The goal of this first mission, says Nozette, is to show that asteroid encounters can be done cheaply, using technologies even more advanced and miniaturized than those that flew on Clementine 1.

NASA's participation in the project is still

to be decided. In its 1996 authorization bill for the space agency, the House of Representatives Science Committee instructed it to use \$5 million of its \$30 million 1996 budget for the New Millennium technology development programme to cooperate on the Clementine 2 mission. But that bill is unlikely to pass this year, making the instruction non-binding.

Project managers for New Millennium say they still intend to cooperate with the Clementine team, perhaps sharing some technology development. But the NASA team has its own plans for an asteroid/comet fly-by mission in 1998, which would also demonstrate a suite of advanced technologies, including solar electric propulsion.

The two missions will in fact be competing fiercely to see which of them can build the cleverest spacecraft at the lowest cost. Competition is good, say science committee members, as long as the two sides do not duplicate each other's efforts.

In both cases, scientists will only be going along for the ride. Ninety per cent of the objectives for the first New Millennium mission are technology-related, even though the programme is designed to produce new sensors and spacecraft capabilities which could then be used on future science missions. And, as with Clementine 1, the Clementine 2 mission will be primarily a hardware demonstration.

This trend towards relegating science to the back seat on demonstration flights worries some space scientists. But Eugene Shoemaker of the Lowell Observatory, who heads an informal Clementine 2 science advisory group that is scheduled to hold its first meeting this week, says that the Clementine 2 team will have more say in selecting science instruments than the first Clementine team did.

Other scientists are concerned that the Pentagon will begin working on strategies to deflect Earth-threatening asteroids before fully understanding the risk they pose. Ironically, it was Shoemaker who chaired a recent panel recommending that NASA fund a long-term observing programme to characterize that risk, but the space agency declined (see *Nature* 377, 185; 1995).

Now, says Shoemaker, the Pentagon may pick up the ball. "NASA didn't bid to do it, and it might then fall by default to the Air Force," he says. Shoemaker and others say it falls naturally within the US Air Force's military mission, and has been seriously discussed in the Pentagon for some time. In fact, says Nozette, taking on the mission of defending against asteroids has generated "surprisingly little opposition in the Air Force hierarchy so far". **Tony Reichhardt**

# Italian science gains budget boost, still awaits reform

**Rome.** Italy's belt-tightening 1996 budget, subject of fierce parliamentary debate since September, was approved by the Senate late last month with some good news for science. Although most other ministries are being required to cut back, the science ministry, headed by Giorgio Salvini, has been spared. Indeed, some research organizations that had been targeted for significant cuts have even been awarded an increase.

The draft budget still has to be approved by Italy's second legislative chamber, the Chamber of Deputies, before the end of this year. But science administrators are optimistic that the Senate's proposals on science will remain untouched.

According to the figures approved by the Senate, the national research council (CNR) will receive a 5 per cent increase next year, raising its budget to IL1,080 billion (US\$679 million). But, according to Sergio Barabaschi, undersecretary of state for science — who, along with Salvini, defended the science ministry's proposals in the Senate through many long nights of acrimonious debate — the victory is greater than the relatively small increase may imply.

Last year, the 1995 budget proposed by the former government of Silvio Berlusconi forecast that the CNR's funds for 1996 and 1997 would be cut to IL850 billion. Furthermore, even this year, the hostility of some members of parliament to the CNR was demonstrated by a proposed amendment requiring all basic research in Italy to be carried out in universities. But the amendment was defeated.

Part of the 5 per cent increase will pay for CNR's planned expansion of research facilities at Monterotondo, near Rome, designed

to host the new European mouse gene repository (see *Nature* 374, 296; 1995). The council plans to move several research groups to the site, where they will work alongside four regional groups of the European Molecular Biology Laboratory (EMBL). The increase is also intended to further the transfer of technology from individual CNR research institutes to industry.

The desire to boost technology transfer also partly accounts for the additional IL50 billion proposed in the draft budget to be added to the current IL500 billion budget of the National Agency for New Technology, Energy and the Environment (ENEA). Last year, the agency found itself under attack from the Berlusconi government over allegations of inefficiency and a lack of focus of its activities (see *Nature* 372, 393; 1994). Its president, the physicist Nicola Cabibbo, has since been concentrating on ways to increase efficiency and accountability of the agency.

Another beneficiary of the increased research budget is the beleaguered Italian Space Agency, ASI, which has been seeking help to meet the substantial overcommitments made during the past decade to the European Space Agency and the US National Aeronautics and Space Administration (NASA), as well as future commitments to the international space station (see *Nature* 373, 467; 1995).

If the budget is approved, ASI funds will increase from IL850 billion to IL950 billion next year, and its budget is forecast to rise to IL1,300 billion by 1998. In contrast, support for the National Institute for Nuclear Physics (INFN) is proposed to remain constant next year, at IL470 billion.

Barabaschi says he is pleased that the Senate has passed a reasonable budget for science in times of economic difficulty and that, for the first time, the science ministry has won some flexibility in handling its budget, allowing it to protect a few smaller research institutes whose budgets have been cut.

According to Barabaschi, the Senate's support of science is also reflected in its approval of a 3.5 per cent increase in this year's IL5,400 billion total research and development budget. But, like many members of the scientific community, he also argues that Italy is not yet ready for the sort of increase in spending proposed in 1993 by the former research minister, Umberto Colombo, namely a steady rise in funding from 1.6 per cent to 2 per cent of gross national product.

Italian research, he says, still has many problems to resolve before it is ready to receive a significant increase in funds, and must first demonstrate that it can spend such funds responsibly.

**Alison Abbott**

# Japan launches plans for Moon probe and new satellite vehicle

**Tokyo.** A more economical satellite launcher and an unmanned Moon probe are among plans set out in a policy paper released last week by the National Space Development Agency of Japan (NASDA), part of the Science and Technology Agency (STA). The release of the NASDA policy paper, the first in six years, has been accompanied by the announcement of a launch date for the first test flight of NASDA's new three-stage solid fuel rocket, the J1.

The new policy paper calls for greater cost-competitiveness in constructing and launching satellites. It describes plans for development of a new launch vehicle, the H2A, which will provide greater lift more



Lift-off: the J1 has its test flight in February.

cheaply than the current H2 launch vehicle. The H2A will be able to lift up to four tonnes into orbit at a cost of ¥8.5 billion; this is less than half the cost of launching the H2, which costs ¥19 billion to launch, and has a maximum payload of only 2 tonnes.

In addition to launching communications, remote-sensing and broadcasting satellites, the H2A, which is scheduled to make its first flight in 2001, will be used to launch the GEM module, Japan's contribution to the international space station.

The new plan calls for the development of an unmanned Moon probe, and for international cooperation on setting up an astronomical observatory on the surface of the Moon. It also suggests that a start should be made on tackling the problem of accumulating man-made garbage in orbit around the Earth.

Although satellite development by NASDA is very advanced technically, it remains commercially uncompetitive, with costs running at more than double those in the United States and Europe. The plan calls for costs to be reduced through subcontracting, using standardized components, and shorter development times.

Low cost and high reliability were the main imperatives behind development of NASDA's new J1 solid fuel rocket, which will make its first test flight on 1 February next year and has been designed to launch small satellites.

**Stephen Barker**

# Stardust mission will collect comet dust

**Washington.** The US National Aeronautics and Space Administration (NASA) has chosen a \$200-million mission to collect samples from a comet as the fourth of its Discovery series of planetary missions. Known as Stardust, the spacecraft will fly past comet Wild-2 in January 2004, photograph its nucleus and trap dust particles with a lightweight aerogel material. The samples will then return to Earth in 2006.

The launch is scheduled for 1999. Principal investigator is Donald Brownlee of the University of Washington in Seattle, and it will be managed by NASA's Jet Propulsion Laboratory. The proposal defeated two other concepts — a solar wind sample return and a Venus atmospheric probe — in competition for Discovery funding.

**T. R.**



# School standards stress scientific literacy

**Washington.** The final version of long-awaited national standards for science education in US schools will be published this week by the National Research Council (NRC), the research arm of the National Academy of Sciences (NAS). Teachers and others hope that the standards will eventually help to raise scientific literacy in the entire population of the United States.

A draft of the standards was published a year ago after considerable controversy. That encouraged schools to stop cramming increasing amounts of scientific knowledge into children's heads, and to emphasize the understanding of scientific concepts, especially through laboratory work (see *Nature* 372, 489; 1994).

After extensive consultation — including the distribution of 40,000 copies of the draft — the final document puts scientific 'knowledge' on an equal footing with scientific 'understanding'. But it argues that such a balance will be achieved only by teaching fewer facts. "The perceived need to include all the topics, vocabulary and information in textbooks is in direct conflict with the central goal of having students learn scientific knowledge with understanding," it says.

The standards include statements of what pupils should know at different stages in their education. For example, it says that, by the age of nine, they "should have a basic understanding of the properties of objects and materials, life cycles of organisms, objects in the sky, and the use of science and technology in solving simple problems".

The standards also set levels of knowledge and ability for teachers — an important issue in a country where many science teachers have had little scientific education themselves. Detailed specifications are offered to local school boards for planning the professional development of their science teachers, as well as assessing the performance of both students and teachers.

The standards suggest that students should be assessed on their ability to apply science to the kind of problems they will encounter in the outside world.

Bruce Alberts, president of the academy and thus chair of the NRC, wants all working scientists to help adopt the report and improve science education in schools. The academy will seek to "mobilize all the scientists and engineers in the country and prepare them to be a force for change in their communities", says Alberts. He suggests interested scientists should spend about four hours a week working with local schools.

Noting that the political tide in the United States is flowing strongly against national standards in general, Alberts adds that the academy will put most of its own effort into promoting the standards at the local and regional level. It will do this through "a year of dialogue" among policy-makers, teachers, scientists, students and parents, which the academy will launch with a meeting in Washington on 9 January and follow up with regional meetings around the country.

The NRC's content standards for 5–17-year-olds are broadly consistent with those published by the American Association for the Advancement of Science (AAAS) in 1993, according to officials of both groups. But the NRC goes further by proposing standards for science teaching, as well as the assessment of such teaching. Alberts says that these initiatives will take ten years to make their full impact on those graduating from American high schools.

Alberts, who arrived at the academy in the summer of 1993, and Richard Klausner, chair of the committee responsible for the project since November 1993 and now the director of the National Cancer Institute, are credited with restoring the project's credibility. It had been underway since 1991, but its initial proposals drew criticism for their alleged reliance on the views of non-

IMAGE  
UNAVAILABLE  
FOR COPYRIGHT  
REASONS

**Science lesson: children should learn fewer facts, but understand more, says the NRC.**

scientists who view science as a "social construct" (see *Nature* 375, 439; 1995).

"No-one at the academy was happy with those early drafts," says Alberts. According to James Trefil, a physicist at George Mason University in Fairfax, Virginia, who sat on a panel which Alberts and Klausner set up to fix the drafts, the project "had fallen into the clutches of post-modernist philosophers". These, he says, saw science as nothing more than "a never-ending change of opinion".

The impact of the standards project will be an important test for Alberts, a prominent molecular biologist who has put education matters squarely at the top of the academy's agenda. Some critics have suggested that, rather than pursuing education issues, he should have been mustering the academy's considerable influence to defend science from threatened budget cuts.

But Alberts is unapologetic. "The good thing about having been a scientist is that I've been working for my whole life on things people say are crazy," he says. "All our [academy] members want us to lobby for every damn thing, but we don't lobby. We're advisors, and you can't be a lobbyist and give independent advice at the same time."

Shorn of the relativist views of post-modernists, and free from the political controversy that has, for example, dogged history teaching standards, the science standards should be able to play a substantial role in revamping science education in US schools, their authors say.

"The scientific community is in agreement on this," says James Rutherford, chief education officer at the AAAS, and a member of a committee that advised Klausner. Trefil adds: "the good news is that there is a consensus amongst the scientists" on what ought to be; the humanities, he points out, are less fortunate.

**Colin Macilwain**

## Champion of medical research bows out

**Washington.** Biomedical research will lose its strongest Republican ally in the Senate next November with the retirement, announced on 1 December, of Senator Mark Hatfield of Oregon. "The whole community is deeply in Hatfield's debt," says Sam Silverstein of Columbia University, past president of the Federation of American Societies for Experimental Biology (FASEB). "He will be very difficult to replace."

David Moore, vice-president of the American Association of Medical Colleges, says Hatfield's loss will be felt not so much in the budget for the National Institutes of Health (NIH), which has support elsewhere, but in the wide range of research issues in which he took an interest.

A five-term senator, Hatfield has been a tireless champion of both the NIH and US medical schools, and took over the chair of the Senate appropriations committee a year ago.

One of Hatfield's legacies will be a seven-year budget bill which includes a reasonable allocation for NIH. In a remarkable testament to Hatfield's influence, 85 of the 100 senators voted in June for his amendment to restore \$7 billion to NIH (see *Nature* 375, 347; 1995). Senator Nancy Kassebaum (Republican, Kansas), the moderate chair of the Senate Labor and Human Resources committee, which oversees NIH, also announced her retirement from the senate two weeks ago.

**C. M.**

# Genome research risks abuse, panel warns

**London.** A panel set up by two US government departments to examine the social implications of the Human Genome Project (HGP) has issued a warning that research funded by the project could be used to support demands for cuts in public funding for social welfare programmes.

The warning comes in a statement from the working group on the ethical, legal and social implications of human genome research — known as the ELSI working group — set up jointly by the National Institutes of Health (NIH) and the Department of Energy, the two agencies jointly responsible for funding the HGP.

The statement is aimed at the arguments raised last year by Richard Herrnstein of Harvard University, and Charles Murray, of the American Enterprise Institute (AEI), in their controversial book *The Bell Curve*, which has already sold more than 400,000 copies. This claims that many social welfare programmes have failed to achieve their objective because they refuse to acknowledge inherent differences in intellectual ability between different social groups.

While distancing themselves from earlier eugenics thinking about the need to eliminate undesirable human traits, the two authors draw heavily on the results of research into the links between inherited characteristics and social behaviour in developing their conclusions.

But in its statement, the ELSI working group claim that genetic arguments cannot and should not be used to inform social policy in the areas cited by Herrnstein and Murray. "Since the lessons of genetics are not deterministic, they do not provide useful information on deciding whether or not to pursue various programmes to enhance the

capabilities of different members of society," says the statement. "Those decisions are moral, social and political ones."

Although *The Bell Curve* was published last year, distribution of the ELSI statement is said to have been delayed by a vigorous debate inside the NIH over how far the agency should become involved in what some claim to be a largely political dispute.

Those responsible for the statement, however, remain convinced of the importance of taking a stand. It has, for example, already been endorsed by the National Society of Genetic Counsellors, many of whose members are concerned that fears of seeing social welfare removed from those considered to have the 'wrong' genetic traits could threaten the social acceptability of therapeutic genetic screening.

"We believe that [*The Bell Curve*] is inappropriately drawing legitimacy from the success of the HGP, misrepresenting the state of genetic knowledge in the area of behavioural genetics, and misusing this information to inform social policy," says Lori B. Andrews, a law professor at the University of Chicago and the chair of the working group.

The statement itself concentrates in particular on the claim by Herrnstein and Murray that high birth rates among the poor, combined with the "dysgenic" behaviour of women with high IQs producing fewer offspring, are threatening the popula-

tion with genetic decline by "exerting downward pressure on the distribution of cognitive ability in the United States".

It points out that research into the complex relationship between genetic and environmental influences on cognitive ability is still evolving and faces many methodological difficulties. To conclude that cognitive ability cannot be changed because it is substantially heritable, and that remedial education is therefore not worth the effort or the cost, "is neither an accurate message from genetics nor a necessary lesson from efforts at remedial education".

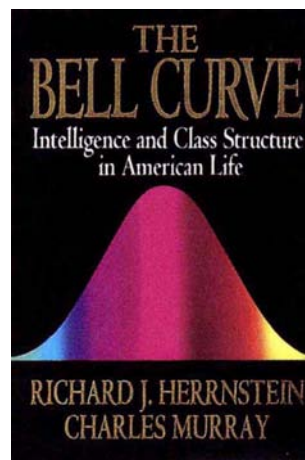
Furthermore, says the working group, a lack of predictability from genetic information is the rule rather than the exception. "Simplistic aims about the inheritance of such a complex trait as cognitive ability are unjustifiable; moreover, as the history of eugenics shows, they are dangerous."

Not surprisingly, the authors have strongly rejected the claims that their conclusions are unscientific — or that they have behaved unethically by discussing the role of genetic influences on social behaviour in general, and cognitive ability in particular.

Although the authors decline to comment on the specific claims made by the ELSI working group, a spokesperson for the AEI dismissed such claims on Monday as being based on "cocktail party discussion" of their book's contents. "The book is more about cycles of poverty than genetics," she said. "It does not have any bearing on the research agenda of the NIH."

Nevertheless, members of the group remain adamant that its concerns need to be taken on board by professional geneticists. Speaking at a meeting at Stanford University at the beginning of last month, for example, Troy Duster, professor of sociology at the University of California, Berkeley, pointed out that the American Society of Human Genetics had declined to sign the working group's statement, claiming that it lacked the expertise to refute the statistical methods used in its calculations.

There was also a warning from David Cox, professor of genetics at Stanford and a key participant in the HGP, as well as a member of the ELSI working group. Discussing the ethical issues raised by the equally controversial Human Genome Diversity Project (see *Nature* 377, 373; 1995), Cox said he was becoming increasingly aware of the extent to which the public tended to see genetic findings as deterministic. "We have a big problem here," he said. □



**The Bell Curve: under attack by the genome ethics group.**

## Medical rivals consider joint operations

**San Francisco.** Stanford University in Stanford, California, and the University of California, San Francisco (UCSF), rivals both on the football field and in the laboratory, may soon collaborate on patient care and medical teaching. Tentative discussions about possible collaboration are the latest in a series of moves by teaching hospitals to adapt to a private health-care system that is becoming increasingly demanding, at a time when federal and state funds are dwindling.

Both institutions have been silent on the specifics of their future relationship. But Joseph B. Martin, chancellor of UCSF, and Gerhard Casper, president of Stanford, said they would focus on economies of scale and preventing duplicated investment in new facilities and equipment. The two medical schools and their faculties would remain independent, with clinical services continuing at both sites.

Eugene Bauer, dean of the Stanford School of Medicine, has told the faculty that a partnership with UCSF would strengthen both institutions. "We have an enormous opportunity, not only to cooperate clinically, but in providing the highest level of medical education in the country." Together, the universities may also increase their ability to cooperate with industry, he said.

Senior faculty and staff at Stanford say they need to focus on improving biomedical research, teaching and clinical care in the face of declining revenues. The faculty may have to shrink while becoming more efficient, they said in a 'vision statement' from a faculty retreat last month.

Casper said that the collaboration aimed to provide a critical mass for teaching in small specialty programmes and more diverse clinical training opportunities at both schools.

**Sally Lehrman**

# Leipzig professor's death stirs bitter feeling

**Munich.** There was a tragic postscript last month to reunified Germany's policy of purging communist staff from universities in the former East Germany when Armin Ermisch, a professor of cell biology at the University of Leipzig who had consistently denied allegations of using his position as a member of the community part to the detriment of his colleagues and students, committed suicide.

Ermisch, who had also strongly denied charges of having ties to the Stasi, died only weeks after he had accepted a financial settlement in return for ending a three-year battle to retain his post at the university.

Along with nearly 1,000 colleagues in the state of Saxony, Ermisch was dismissed from his post in 1992 after a brief hearing by a committee, known as a *Personalkommission*, appointed by the science minister for the *Land*, Hans-Joachim Meyer. Using information from Stasi files and university colleagues, the committee decided Ermisch had abused his communist party links during the twenty years before the fall of the Berlin Wall (see *Nature* 359, 762; 1992).

After reunification, all university staff were required to submit themselves to examination by such *Personalkommissionen*, normally made up of public figures, academic colleagues and students, as well as one or two professors from west Germany. The fairness of these committees has frequently been challenged, many critics arguing that the lack of checks allowed personal grievances to influence their recommendations.

The final decision on all dismissals rests

with Meyer personally. Ermisch appealed against his own dismissal and, in the summer of 1993, won his case when a court found no evidence that he had abused his communist party links, and said that there were no grounds for dismissal.

But Ermisch did not get his job back. In the period after his dismissal, the entire university structure had been overhauled in line with reunification laws. Faculties were reorganized and streamlined, and all academic staff were required to reapply for their jobs in open competition. Staff numbers



**Ermisch: remembered as a brilliant teacher.**

were limited, and universities were entitled to dismiss superfluous staff within the restructured system.

Determined to complete the reunification process quickly, Meyer chose not to wait for the outcome of lengthy and numerous court appeals against individual dismissals, and during this process Ermisch's post had been filled. Ermisch appealed for reinstatement, and earlier this year, the court once again ruled in his favour, instructing the ministry to reemploy him. But the science ministry countered again with the argument that all vacancies had been filled.

According to Rainer Landgraf, a neurobiologist who was trained by Ermisch and now

works at the Max Planck Institute for Psychiatry in Munich, Ermisch became very depressed when he finally accepted that he would never work again, and was forced to reach an out-of-court settlement. Landgraf has always supported Ermisch in his fight for justice. "Everyone remembers him as a great scientist and brilliant teacher," he says.

John Russell, head of the department of physiology at the University of Edinburgh, and a scientific collaborator of Ermisch, points out that Ermisch's frankness and honesty may itself have stimulated opposition within his faculty. Landgraf says that, as head of department, Ermisch never obstructed the careers of anti-communist colleagues, and judged them — albeit strictly — only on scientific grounds.

Another of Ermisch's supporters is Wolf-Dietrich Arnold, head of the University of Leipzig's orthopaedic clinic, who, like Ermisch, successfully defended himself against charges of abuse of power. "The way Saxony has worked has been particularly harsh," says Arnold.

Candidates shortlisted for Arnold's job gave lectures in his department the day before the results of his first, successful, appeal were announced. After winning his second appeal, Arnold decided to abandon his fight and moved to a non-academic post in neighbouring Thüringen. He says that he regrets no longer being able to carry out research, but that "what hurts most is losing contact with students".

Ermisch's case is not an isolated one. More than 300 professors appealed against their dismissals on the basis of *Personalkommissionen* recommendations. Of the 250 cases that have reached conclusion, most, like Ermisch, finally accepted settlements. Twenty-five have won their appeals, but only a third of these have been given their jobs back.

The GEW, the trade union that represents German teachers, describes the situation as hopeless and unfair, pointing out that the newly appointed staff also now have employment rights. A GEW spokeswoman says that the state of Saxony must now work out how to obey court orders to reemploy academics judged to have been unfairly dismissed on the basis of *Personalkommissionen* recommendations. A major problem, she says, is that the state has been able to restrict itself to individual cases, as there is no collective movement to put pressure on government authorities.

The science ministry remains quiet on the issue, merely repeating that it has always followed the letter of the law, and that, if individual mistakes have been made, they do not outweigh the greater benefit of having quickly removed the powerful influence of communism from east German universities.

**Alison Abbott**

## AIDS survey traces shifting attitudes

**Paris.** French people are becoming increasingly aware of the risk of HIV transmission and changing their sexual behaviour as a result, according to a national survey on attitudes to AIDS.

The survey, which was conducted in November and December last year and is based on a random sample of 1501 individuals, confirms that the use of contraceptives is continuing to increase, particularly among the under-30s and single people. But more than a quarter of those people who have sexual relations with several partners say that they sometimes have unprotected sex.

Responses among the French population to the spread of HIV/AIDS involve a range of factors, however, in addition to the use of contraceptives. Some of those cited by respondents include screening tests by individuals and their partners, care over the selection of partners, and a reduction in the number of casual sexual partners, particularly among those over 30 years old and living as a couple.

This last factor marks a change when

compared to previous surveys, and the finding remains to be confirmed, according to Nathalie Bajos of the French national institute for medical research (INSERM). Further analysis should reveal whether this reduction in the number of casual partners is the result of bias in the responses of interviewees, or whether extra-marital relationships are perhaps becoming less acceptable to society.

The survey also reveals an ambiguous attitude towards those who are HIV-positive. While public tolerance towards infected individuals is growing, an increasing number of people favour coercive measures, such as obligatory screening or the creation of special medical centres to house them.

More than 80 per cent of those surveyed say that they are prepared to work, eat and go on holiday with someone who is seropositive. But almost the same number, 78.9 per cent, would like to see specialised medical establishments set up to take responsibility for patients infected with HIV.

**Catherine Tastemain**

# UK science budget avoids the spending axe

London. Britain's scientific community last week gave a collective sigh of relief as the government announced that it had decided to hold constant the £1.3 billion annual science budget in real terms for 1995-96, despite pressure to reduce public spending to help finance a one per cent cut in the basic rate of income tax.

But there remains concern that much of the extra £30 million being awarded to the research councils to keep up with inflation will be swallowed up by increased subscriptions to international facilities — in particular the European Laboratory for Particle Physics (CERN) and the European Space Agency — due to the falling value of the pound.

John Mulvey, of the pressure group Save British Science, said that he congratulated Ian Lang, president of the Board of Trade, who has cabinet responsibility for science, for protecting the science budget. "We ought to give him credit for a good job," says Mulvey, one of a number of critics to suggest that last July's transfer of the Office of Science and Technology (OST) to the Board of Trade amounted to a step down

for science. Lang has "done well to retain science spending", he concedes.

Richard Brook, chief executive of the Engineering and Physical Sciences Research Council, said that scientists could take "broad comfort" from the fact that funding levels had been held up. "We should credit the OST with having held the government's loyalty to science through what everyone agrees has been a difficult budget," he said last week.

News of the science budget was also welcomed by Sir Michael Atiyah, outgoing president of the Royal Society, which receives a share of the budget along with the research councils, the Royal Academy of Engineering and the special initiatives that are overseen by the OST. At the same time, however, both Mulvey and Atiyah point out that the government's budget statement included projected annual cuts to science spending of around £14 million over the following three years.

Mulvey, in addition, says he is concerned at the fall in capital funding for universities, part of the education budget, by £300 million in real terms. "The science base

stands on two legs: research councils and the universities higher education funding council," says Mulvey. "We now have a situation where one leg has done well, while the other has been hacked off."

Kenneth Pounds, chief executive of the Particle Physics and Astronomy Research Council (PPARC), says he "shares the relief of other research councils" that the budget has been held steady. But he remains concerned about how much of PPARC's budget will be used to pay Britain's subscription to CERN and ESA. "We can be certain it will go up," he says, although refusing to confirm rumours that additional costs may be as high as £20 million.

As a result of changes introduced as part of the White Paper of 1993, this extra cost will now be shared by all the research councils, although there is still discussion taking place as to the size of the share that PPARC should shoulder during the current transitional period. A statement on the CERN subscription is expected this week.

Allocations for the science budget between the six research councils will be announced early in 1996. Ehsan Masood

## Rabbit virus threatens ecology after leaping the fence

**Sydney.** The Australian government is considering whether to bow to the inevitable and officially release a rabbit virus that escaped from its test site on a remote island in late October. But some Australian scientists are predicting a wave of extinctions following the breakout of rabbit calicivirus disease (RCD) virus, two years before it was due to be officially released as a biological control.

Few in Australia have been sorry to see the deaths of at least one million rabbits by the end of November, as rabbits are regarded as vermin. But they had also become natural prey for other animals, such as foxes and feral cats.

State and federal departments of agriculture have started culling those predators. But Glenn Albrecht, a lecturer at the University of Newcastle's geography department, and a vocal critic of the RCD programme, says that the *ad hoc* and rushed culling programme will not save many species. "There is potential for major ecological damage to Australia," he says, in part "through a mass wave of extinctions due to prey switching".

RCD is a naturally occurring virus. It killed millions of rabbits in China and Europe during the 1980s, without seeming to harm other species, and was therefore an obvious candidate for use as a biological control in Australia.

Apart from competing with native animals for food and destroying native vegeta-

tion, rabbits cost farmers an estimated A\$120 million (US\$89 million) a year.

One attempt has been made to eradicate rabbits with the mosquito-borne disease myxomatosis, which was also prematurely released in the 1950s. But although the disease still works to a certain extent, the rabbit population has built up an immunity to it and is almost back to its pre-myxomatosis levels.

After extensively testing RCD virus on a range of native animals and finding it had no effect, the RCD committee, which includes scientists from a number of organizations, was testing its effects on rabbit populations on Wardang Island, off the coast of South Australia. A spokesman for the RCD com-

mittee said it is still did not know how the virus escaped from the island. The best guess was that it was spread by bush flies (also very common in Australia), aided by warm weather and a strong southerly wind.

He said the virus now appeared to be spreading north and west but not towards the east, after initially getting as far as the mining town of Broken Hill, over the state border in New South Wales. He said that the pattern of spread appeared to be related to rabbit population density.

Faced with the reality of the premature release of RCD, the Biological Control Authority, which reports to a joint committee of state and federal government ministers responsible for agriculture and resource management, has called for submissions on whether RCD should be officially released. It will also decide whether to recommend an official inquiry into the reason for the outbreak.

Whatever its cause, the premature release has led Australian scientists and conservationists to question whether scientific procedures for testing viruses are at all adequate, and even the need for biological controls.

Stuart Pearson, who along with Albrecht is also a lecturer at the University of Newcastle's department of geography, says that there is still a small risk of Australian mammals being infected by RCD, as not all have been tested for their reaction to the virus.

Mark Lawson





# Patents and indigenous lore

SIR — Commenting on the controversy about patenting products from the neem tree (*Nature* 377, 89; 1995) you gave the opinion that while we cannot dismiss the claim of unfair discrimination against indigenous knowledge, the patents should be judged on (their genuine originality and) the terms on which they were granted, not the extent to which they conflict with *traditional knowledge systems that have only a marginal role in the modern world* (my emphasis). The trouble is that today's patenting system permits companies and individuals to lay claims to techniques based on the indigenous knowledge of developing countries, if they are able to extend this knowledge using modern methods; in other words, a veneer of modernism over traditional 'prior art'. Traditional knowledge systems play more than a marginal role in the modern world; indeed, they form the basis of the 'drug hunting' in which companies and inventors are involved, whether in Costa Rica or Southern Asia. How did W. R. Grace come to know of the properties of neem?

To that extent, patenting extraction procedures or products from sources that were well-known, albeit through empirical or traditional wisdom, would amount to 'folk wisdom piracy', because it turns public good into a private commodity. The coalition of 200 Green groups that is petitioning against the W. R. Grace-USDA patent urges us not to forget that such commercial interest in the neem tree will lead to large-scale purchase of neem seeds at US\$300 per tonne, essentially pricing these seeds beyond the reach of local farmers and forcing them to rely on the modern formulated product rather than traditional recipes.

**D. Balasubramanian**

*Centre for Cellular & Molecular Biology, Hyderabad 500 007, India*

SIR — India's remarkable neem tree grows wild wherever it is allowed to in the area in which I live (13.04° north of the Equator). Indian women have for centuries placed sprays of the neem among stored clothes to deter insect pests, and the neem's insecticidal properties have always been widely recognized.

Perhaps what is less known is its purported medicinal value. When smallpox was still rampant, poultices of neem leaves were used to relieve the unbearable burning of the lesions. Indian villagers may be excused for believing that the neem will also alleviate other clinical conditions. Thus at two villages, one not far north of Madras, an annual fair is held when pilgrims first bathe and then parade around a small shrine dedicated to Mariamma, believed to be the goddess of smallpox, wearing nothing but small garments of neem leaves. Far from vanishing with the disappearance of smallpox, this annual festival has been gaining in popularity over the years and is an amazing spectacle.

But if the neem is so efficient an insecticide, why is the tree itself so prone to a devastating insect attack almost every year? Just before the very hot weather begins in this part of India, all our neem trees are attacked by an insect, with the result that the end leaves of the trees turn a dirty brown and fall off (see photograph).

One year this infestation was so bad that owners of affected neem trees were preparing to cut them down in the belief that the trees had died. But the trees recovered as they always do.

The Chief Conservator of Forests of the State of Tamil told me that the insect responsible is the 'tea mosquito', which does also attack tea as well as cashew and other plants. No one, however, seems to

know the taxonomic name of this parasite or any other details about it.

The mystery, therefore, is why a tree renowned for its insecticidal properties is itself so vulnerable to the attacks of an insect.

**Harry Miller**

*3A Satyanarayana Avenue,  
Boat Club Road,  
Madras,  
India 600 028*

## Anticancer drugs

SIR — Petsko<sup>1</sup> (rightly) congratulates Takahara and colleagues for obtaining the high-resolution crystal structure of the cisplatin-DNA adduct<sup>2</sup>. He also suggests that related heavy-metal compounds may have antitumour properties, and that such complexes may be "attractive candidates for rational design projects". But is the pursuit of new anticancer drugs that non-specifically damage DNA worth large expenditures of time and money?

Despite the introduction of numerous chemotherapeutic agents, cancer survival rates for clinically advanced tumours have improved little in the past 30 years. With the exception of testicular cancer and a few other rare neoplasms, solid malignancies when disseminated remain incurable, and treatment can only be palliative. Failure of treatment with chemotherapy is due generally to *de novo* intrinsic resistance of subpopulations of tumour cells to anticancer drugs, and, after several cycles of therapy, resistant clones predominate in the tumour cell population<sup>3</sup>. Thus the physician encounters the common clinical scenario in which a cancer initially regresses after chemotherapy, followed by lack of tumour response to further drug treatment and, invariably, progression of the disease. It is unlikely that new agents (with similar mechanisms of action) will circumvent such resistance.

Clinical and molecular oncologists must acknowledge that nonspecific cytotoxic chemotherapy drugs will never be effective in the treatment of the common solid cancers. The continued development of these agents will result at best in marginal improvements in survival and quality of life, and will require the use of enormous resources. Such resources would be better used further to define the molecular alterations that cause cells to become malignant, and to design therapeutics that target those alterations.

**David M. Reese**

*Division of Hematology/Oncology,  
UCLA School of Medicine,  
Los Angeles,  
California 90095, USA*

1. Petsko, G.A. *Nature* 377, 580-581 (1995).

2. Takahara, P.M. et al. *Nature* 377, 649-652 (1995).

3. Gottesman, M. M. & Pastan, I. *Trends pharmac. Sci.* 9, 54 (1988).



An old neem tree in Madras after attack by the 'tea mosquito'.

# Science in an occupied country

SIR — Latvia is an independent country on the eastern shores of the Baltic Sea that was occupied by the Soviet Union in 1940. On 16 June that year, the president of Latvia received an ultimatum to change the government and to permit the Red Army to cross the Latvian borders. The next day, Soviet tanks filled the streets of towns and villages of Latvia. The government retreated. The invaders appointed a new one.

The attitude of the Soviet occupiers towards Latvian intellectuals is typified by the fate of Professor Jūlijs Auškāps, the minister of education from 1938 to 1940, who had formerly been rector of the university. Auškāps had been active in the organization of science. Immediately following the occupation he was dismissed, forbidden to teach at the university and even to live in the capital, Riga. In June 1941 he was deported and imprisoned, and in August 1942 he was executed in Sverdlovsk. His 'crime' was being a Latvian patriot. The next minister of education in Latvia, Jūlijs Lācis, a popular novelist, although appointed by the occupiers, was arrested in January 1941 and died in prison in Astrakhan in December that year. His successor was an insignificant man from Moscow.

Science in independent Latvia in the 1920s and 1930s was concentrated in the newly founded Latvian University in Riga where professors had good links with colleagues in other European universities. On the day of the occupation, the Iron Curtain fell, and direct communication with the West, as well as the flow of chemicals and equipment, stopped. The activities of the university were restricted and the faculties of theology and philosophy were closed. The deans of six faculties lost their posts on 14 June 1941 when 15,000 Latvian citizens were deported overnight to Siberia in locked cattle-wagons<sup>1</sup>. Among those deported were ten professors and more than a hundred students of the Latvian University.

In July 1941, the Soviet occupation of Latvia was succeeded by German occupation. The Latvian University was not restored and there was only limited activity at the "University in Riga". All Jewish professors and students were arrested, imprisoned in the Riga ghetto and executed. Many students were conscripted into the German army. When that army retreated in 1944 and a second Soviet occupation became inevitable, most of the remaining professors fled into exile in Germany and Sweden. They were afraid of deportation by the Soviets and other forms of genocide against Latvians.

In March 1949 there was another mass-deportation of Latvians to Siberia. This time the locked cattle-wagons carried 42,000 people. Hundreds of talented students and scientists of the new Academy of

Sciences were lost. They were allowed to return to Latvia only after 1956, but many died from cold and hunger in the inhumane conditions of Siberian exile.

The invention of the atomic bomb at the end of the Second World War increased the prestige of science. That was one of the main reasons why, at the end of the 1940s, by command from Moscow, academies of sciences were founded in all the Soviet republics. The Latvian academy was founded in 1946; in 1990, it included 20 separate institutes employing 4,668 people, of whom 1,059 held at least a doctorate. During that period we made some successful investigations especially in physics and chemistry. But some of those entering science did not have high intellectual abilities or were not sufficiently interested in science — and much of the laboratory equipment was outdated. Restrictions were imposed on us. All investigations about marshlands, even botanical, were secret. The Soviet governors believed that marshlands were primarily an obstacle to enemy armies and should be kept secret. Facts about industrial pollution of land and water were not published. Similarly, all data about epidemic diseases were secret.

After restoration of the independence of the Republic of Latvia (Parliamentary Declaration of Independence on 4 May 1990) there began a period of transition of the economic system from "socialism" to a "free market". But the government of Latvia can afford to spend only 0.35 per cent of gross national product on science. In these conditions, there is a fast deterioration of science that can be characterized by the number of investigators with a doctorate working in the institutes of the Latvian Academy of Sciences: 1,059 in 1990, 854 in 1992 and 696 in 1995. The salary of an assistant is about US\$80 a month and that of a professor US\$180.

The remaining scientists are trying to keep up their work in the hope of better times. In Latvia, as well as in Lithuania and in Estonia, we receive aid from some foreign foundations, notably the Soros Foundation. Some of our young scientists are able to study at universities and to work in laboratories in Western Europe and in the United States. Meanwhile, we are seeking a new platform for the science of the formerly occupied Latvia — incorporation into European science<sup>2</sup>.

**Pēteris Ozoliņš**

*Department of Physiology,  
Latvian Institute of Experimental  
and Clinical Medicine,  
O. Vācieša ielā 4,  
Rīga LV-1004, Latvia*

1. *These Names Accuse: Nominal list of Latvians deported to Soviet Russia in 1940–41*, 2nd edn (Latvian National Foundation, Stockholm, 1982).
2. Dickson, D., Abbott, A. & Levitin, C. *Nature* **372**, 591–610 (1994).

# Restoring good manners

@SIR — Simon Wolff (*Nature* **377**,192;1995) attacks peer review as practised by journals and funding agencies. We cannot speak for journals, but can comment from the perspective of a major British medical research charity that has always taken peer review very seriously and continually reviews its procedures. We have also recently contributed to reviews of the UK peer review process with both the Association of Medical Research Charities and the Royal Society.

Wolff says that there is "total collapse of peer review within journals and by grant-making organizations". This misrepresents the system, exaggerates the difficulties and unnecessarily questions the integrity of organizations that are an essential part of the scientific community and, in our view, strive to serve it well. We are constantly amazed and humbled by the trouble reviewers take on behalf of the community in general and the Wellcome Trust in particular: rarely are their efforts given recognition or thanks. Reviewers devote long hours to reading the proposals, supporting documents and often the relevant literature, both extant and as pre-prints. Their reports are frequently long and not infrequently as detailed as the application; they are rarely other than honourable in intent and objectively critical in content. We do not recognize the "bad-tempered nature of science", at least as it concerns our activities.

The Wellcome Trust, for its part, goes to considerable lengths to consult a large body of reviewers, both from the United Kingdom and overseas, using both the personal knowledge of our expert staff and advisers and electronic databases and retrieval systems. In the grant year 1994–95, the Wellcome Trust used more than 3,600 reviewers on more than 2,100 grants, of whom 34 per cent were based overseas. We ensure that reviewers are not over-used by checking in our database precisely how frequently individuals are asked to review for us. We do not believe that the process is generally regarded, or ever should be regarded, as nothing more than a marketing exercise. It is, warts and all, generally accepted as the best method to evaluate scientific research.

Neither the peer-review process nor the Wellcome Trust or other grant-funding agencies are perfect: they do not however deserve to be traduced.

**Michael J. Morgan**

*(Chairman, Science Funding)  
The Wellcome Trust,  
183 Euston Road,  
London NW1 2BE, UK*

■ Sadly, Simon Wolff died on 25 November. □

# The case for genomic patenting

George Poste

**Opposition to the patenting of genomic inventions threatens to erode the foundation of intellectual property rights needed to convert innovative research into new drugs, vaccines and diagnostic tests.**

THE perceived importance of genomics in creating novel drugs, vaccines and diagnostic tools is reflected in the accelerating pace of patent applications for genomic inventions filed by academic institutions and industry. The largest databases of information on human gene sequences are believed now to be in industry. But the distinction between academic and industrial research in genomics is increasingly ill-defined. Academic researchers have moved beyond their traditional role of industrial consultation to help establish new companies in which they, and their institutions, may hold financial interests and assign intellectual property rights to inventions made in their laboratories using taxpayer-sponsored government grants. These trends have evoked heated debate about the intellectual and moral legitimacy of genomic patents and their effect on the progress of research<sup>1-4</sup>. Here I present the case for genomic patents, challenge the claim that they impede research and examine the implications of proposals that human gene sequences should not be patented but placed in the public domain.

The applicability of intellectual property laws to genomic inventions has largely been relegated to a secondary aspect of the debate (for a review see ref. 5). For many participants, the patents controversy offers a banner of convenience under which to examine issues far removed from intellectual property law. The academic community primarily worries about whether patenting hinders the free exchange of knowledge and whether industry is usurping patents unfairly without recognition of academic contributions<sup>6,7</sup>. For religious and anti-technology lobbies, opposition is part of a broader concern about the moral, ethical and social implications of genetic research and the moral boundaries of science. For industrial participants, the motivation, unsurprisingly, is the quest for commercial leadership in genetic medicine.

## Principles

Although patent laws in Europe and the United States differ, they share the principle that patents are granted only for inventions that are novel (not previously disclosed), inventive (must not be obvious) and have utility (practical use as opposed to purely intellectual or aesthetic merits). Opponents of genetic patents have argued that identification of a gene (or a gene subfragment) is a discovery and not a patentable invention. But the skills

required to construct full-length genes and define their function and utility are not straightforward<sup>7</sup>. They have been accepted by patent authorities in Europe, the United States and Japan as evidence of inventiveness, with a rapid growth in the number of gene patents awarded.

For some in the research community, academic and industrial, concern stems less from patenting *per se* than from the prospect that patents will be awarded for the mere sequencing of genes or, worse still, gene fragments, without elucidation of function and utility. This debate was sparked in 1992 when the US National Institutes of Health (NIH) decided to seek 'omnibus' patents for several thousand complementary DNA fragments, or expression-sequence tags (ESTs), of unknown function<sup>5</sup>. Following widespread condemnation from academic institutions, industry and foreign governments, the NIH withdrew the patent file. The intellectual property status of ESTs remains ambiguous. But the central issue in the EST controversy applies equally to full-length genes: should patents be awarded for any genome sequence without evidence of its function and utility?

The patenting of genomic inventions, provided that function and utility are demonstrated, has been endorsed by diverse academic, industrial and governmental groups<sup>7-10</sup> including the Human Genome Organization (HUGO), the Biotechnology Industry Organisation, the UK Medical Research Council, INSERM in France, the Danish Council of Ethics, the UK Nuffield Council on Bioethics, the NIH, the UK Patent Office and the UK House of Commons Science and Technology Committee, and in testimony to this committee by the four leading pharmaceutical companies in the United Kingdom: Glaxo-Wellcome, Zeneca, Pfizer and SmithKline Beecham<sup>8</sup>. Assuming that novelty, inventiveness and utility are demonstrated, ESTs should also be patentable. But in reality their utility is likely to be more limited than that of full-length genes. ESTs have commercial value as diagnostic tests of altered gene expression in disease and as structural templates for oligonucleotide-based therapeutic agents such as ribozymes and antisense and triplex molecules. These applications meet the criteria for patent protection. That the function of each EST is not defined, and may not be for many years, does not detract from the validity of claims for novelty (detection of

previously unrecognized differential gene expression in disease) and utility (a diagnostic fingerprint for disease detection or staging). Granting such patents is no different from the patenting of other biomarkers such as the *BRCA-1* breast cancer gene. Its function in physiology or pathology remains unknown, but its use in testing for predisposition to disease is indisputable.

## Equity

HUGO has expressed concern over how intellectual property rights can "be allocated fairly, in a manner that appropriately weighs the contributions of different parties to the total research effort, and creates necessary incentives for the ongoing development of products without interfering unduly with scientific research"<sup>7</sup>. Increasingly, many research groups create the technical foundation to allow the final step in gene identification and characterization, but the patent accrues only to those who solve the 'last step'. There is no equitable solution under current law, other than for competing parties to negotiate co-inventorship, as was done by the NIH in asking Myriad Genetics to add US government scientists as co-inventors in the patent claim for *BRCA-1*. Inventorship is a matter strictly of law, however, not negotiation by convenience.

Prohibition of patenting of human genes is advocated on grounds ranging from theological claims for the sanctity of human DNA to claimed moral imperatives to protect the human genome from commercial avarice and to safeguard society against perceived discriminatory and eugenic threats from modern genetics. Several categories of 'natural' materials are exempt from patenting but no special status is yet accorded to DNA, RNA, synthetic oligonucleotides or the products they encode<sup>11</sup>. Patent laws treat them straightforwardly as compositions of matter or articles of manufacture<sup>11</sup>. The courts have ruled consistently that nucleotide sequences when removed from the body are patentable<sup>5,11</sup>. But European and US patent laws differ: the former contain specific provisions to deny patenting on the grounds of moral offensiveness or threat to public order<sup>11</sup>. Patent examiners have been reluctant to invoke this provision, for several reasons. First, the morality provision applies not to the invention but to its use (that is, exploitation). Second, if use of an invention is judged sufficiently abhorrent to be deterred, alternative action to make it

illegal is a far more powerful constraint. Third, public morality has ever-shifting boundaries: who has the wisdom or foresight to say how society will define 'moral' over the full 20 years of the patent life?

It is appropriate for society to scrutinize the direction and moral dimension of science and its applications. But it is neither appropriate nor productive to expect patent law to resolve issues it was never designed to address. If society sees a compelling moral, religious or social need to modify the conduct of scientific research, it should be pursued through other forums of public debate more suited to setting legislative guidelines and statute.

Recent calls for compulsory licensing of patents for detection of hepatitis C virus in blood and the use of the *BRCA-1* gene in breast-cancer screening argue that some genes are of sufficient public benefit that commercial monopoly should be prohibited. Although playing well to the public gallery and hence often endorsed by politicians, such claims ignore legal and practical realities. What criteria would be used to define the scope of the public interest and who would define it? Who would be responsible for ensuring that the claimed public benefit is realized by rapid commercialization? Who would be held accountable if the inventions languished without commercial sponsorship? How would differing 'national interests' be harmonized internationally? Notwithstanding these complexities, placing inventions in the public domain typically results in delay or overt abandonment of commercialization<sup>12,13</sup>.

## Incentives

Industry has long held that intellectual property rights are major incentives for the investment in research and development (R&D) needed to commercialize new drugs<sup>12-15</sup>. Securing intellectual property on publicly funded academic research, and the granting of exclusive rights to industry, are recognized by government directives on science policy in the United Kingdom<sup>12-14</sup> and the United States<sup>16</sup> as important incentives to aid technology transfer from academic institutions to industry. Patents enhance competitiveness by forcing companies to adopt new research strategies and explore new disease targets, thereby catalysing breadth and depth in research innovation. Patent protection is also valued by the financial and pharmaceutical sectors when evaluating investment opportunities in new entrepreneurial companies.

The recent UK parliamentary report on human genetics expressed concern that "patent examiners are applying the criteria of utility and novelty too loosely" and recommended that the European Patent Convention be "redrawn to allow patents to be challenged on grounds that their claims go too wide"<sup>18</sup>. Granting all-embracing claims rewards patent-holders for future inventions they have neither

demonstrated experimentally nor, worse still, even envisaged. Unsubstantiated breadth of claims is alien to the original intent of patent law. It confers commercial monopoly without disclosure, deters competition and erodes the credibility of the patent system. Patents are intended to reward specific inventions and to stimulate competitiveness by forcing others to explore new avenues of research while protecting the patent-protected field of enquiry from competition. Dilution of the integrity of the patent system, whether through undeserved competitive dominance from overly broad protection or through the ambiguity produced by overlapping claims on the same invention and the increasing need for extensive cross-licensing, can only degrade the value of intellectual property.

Some academic commentators argue that patenting genes, proteins and other materials impedes the progress of basic research by hindering the free exchange of knowledge and research reagents. Yet academic institutions now rank with industry in the intense pursuit of patents on molecules used by the basic research community<sup>17</sup>. In any case, the generosity of academic researchers in exchanging reagents or new findings has always been variable<sup>18</sup>. Some share them openly, whereas others withhold them or set restrictive conditions to hinder exchange. But most universities now impose a material transfer agreement on reagents sent to other researchers, academic and industrial. These grant the donor university or investigator legal rights to any invention made by the recipient laboratory using the transferred materials<sup>6,19</sup>. The recent NIH 'universal' material transfer agreement places great emphasis on the protection of the donor's intellectual property rights<sup>19</sup>.

That these trends are changing the environment for academic research is not disputed. Whether they impede basic research that has no immediate commercial application can probably be assessed only case by case<sup>6</sup>. Although some may view them as alien to academic scholarship and cooperation, the changes are driven from within academic institutions without intervention from industry. They arise from the new emphasis in government policy to get value from taxpayer-sponsored research and the greater importance universities and academics now put on intellectual property rights as a possible source of financial gain<sup>20</sup>.

Finally, patent law already provides a broad exemption for the use of proprietary materials for 'research purposes'<sup>5</sup>. It serves little purpose, from public-relations or financial standpoints, for industry to prosecute academic institutions for infringing a patent while carrying out basic research. If, however, the university, or its researchers, seeks to commercialize an invention made while infringing an existing patent, the litigation risk would increase.

A ban on patenting genomic inventions would invariably mean that companies, or any party, wishing to use the knowledge to create products would resort to trade secrecy. This would slow research, with inventors, academic and industrial, withholding knowledge to the detriment of the overall research community. It would affect adversely the culture of industrial R&D laboratories and hinder collaboration between academic and industrial sectors. By forcing industrial researchers and academic collaborators to withhold publication or public presentation, secrecy would create a destructive, anti-intellectual climate at a time when ties between academic institutions and industry in the life sciences are prospering and industry is increasingly able to attract gifted researchers because of enlightened, open-publication policies.

Genetics benefits from transparency and open debate. A forced retreat to secrecy would lessen public insight and predispose the research community to constant challenge by the misinformed, the mischievous and the malicious. Those seeking to pioneer genetic medicine might usefully reflect on the fate of the nuclear power industry and the consequences of failure to educate and inform a concerned public.

## Access

The largest repository of human genetic sequence data is believed to reside in industry, with three databases predominating: (1) a coalition of The Institute for Genome Research (TIGR), Human Genome Sciences (HGS) and SmithKline Beecham (SB); (2) a separate HGS database; and (3) InCyte Pharmaceuticals. Only the first is currently available for use by researchers from any non-profit organization<sup>21</sup>, including academics with financial interests as company officers in for-profit entrepreneurial companies. The participating university agrees to notify the database sponsor of any patentable invention made with the data or clones received. But the intellectual property right remains with the academic institution. No obligate commercial rights are granted to TIGR-HGS-SB. These rights must be negotiated case by case. If they are not negotiated within six months, the institution is free to negotiate with others but cannot offer terms less demanding than those offered to the database owners. With merely an *option* to negotiate commercial rights, and intellectual property rights remaining with the institution, the TIGR-HGS-SB database terms are no more onerous, and in many cases less so, than the material transfer agreements between universities and other academic centres or industry. Indeed, the NIH 'universal' material transfer agreement seeks to grant the donor *automatic* rights to any invention by a third party<sup>19</sup>.

Genomics is imposing a new competitiveness on industry, stemming from differences in the views and strategies of



companies. Merck claims that in most cases genes and ESTs are 'research tools' that should be placed in the public domain and should not be subject to patenting<sup>22</sup>. The company has funded Washington University in St Louis, Missouri, to establish a public database of human cDNAs to "accelerate the gene discovery process" and to "benefit human health"<sup>22</sup>. The biotechnology industry and many pharmaceutical companies, including SmithKline Beecham, consider that some genes and proteins also have potential as drugs and diagnostics in their own right and thus seek to protect them by patents. By contrast, Merck considers that direct therapeutic use of genes and proteins will be "rare"<sup>22</sup> and that "for the vast majority of expressed genes the principal value will be as a research tool"<sup>22</sup>. By viewing genes and other macromolecules as tools and targets for drug action, rather than as drugs, Merck can afford to donate them to the public domain and instead seek patent protection on synthetic molecules that act on 'target' genes or proteins. If public donation of sequence data on expressed genes meanwhile dilutes their value to competitors, this act of seeming altruism will create adverse circumstances for competitors.

The existence in the public domain of the nucleotide sequence of a portion of a gene, or an extended amino-acid sequence of part of a protein, may constitute a level of 'obviousness' or 'anticipation' sufficient to deny a patent to those who subsequently discover the full gene or the complete protein and elucidate function<sup>5</sup>, although this may not apply in all cases<sup>23</sup>. Any group, academic or industrial, embarking on a campaign to place extensive human EST gene sequence data in the public domain could create a major obstacle to others seeking to patent the full-length gene and its function at a later date. Such actions, even when in the 'public interest', could reduce a company's prospects of securing intellectual property rights. For small companies founded entirely on the use of one gene or protein, this threat is profound.

In advocating that research tools should be placed in the public domain, Merck has

#### RECENT PATENTS FILED BY MERCK, SHARPE AND DOHME FOR RESEARCH TOOLS: TRANSGENIC ANIMALS, SCREENING ASSAYS AND COMPLEMENTARY DNA SEQUENCES

|                |                                                                                         |
|----------------|-----------------------------------------------------------------------------------------|
| WO 9503397     | Transgenic animal model for cognitive disorders                                         |
| WO 9503402     | Transgenic animal model for human interleukin 1-β                                       |
| WO 9421784     | Stromelysin-1-deficient transgenic mouse                                                |
| GB 282814      | cDNA for E2F-2 transcription factor                                                     |
| WO 9509872     | cDNA for prostaglandin receptor IP for screening for receptor modulation                |
| WO 9512661     | Coding sequence for triol polyketide synthase                                           |
| WO 9510625     | Coding sequence for 1,3-β-D glucan biosynthetic enzymes                                 |
| WO 9504821     | Coding sequence for glucagon-like 1 peptide                                             |
| WO 9517416     | Coding sequence for Wnt-x growth factor                                                 |
| WO 9517516     | Mouse genome locus for high-level recombination of genes                                |
| WO 9503065     | Purified scorpion venom peptide margatoxin                                              |
| WO 9421660     | cDNA encoding α1-1C and -1A human adrenergic receptor                                   |
| WO 9410184     | cDNA encoding mammalian farnesyl protein transferase                                    |
| WO 9414977     | cDNA encoding human cyclooxygenase                                                      |
| WO 9413799     | Stable transfection of eukaryotic cell line expressing human GABA <sub>A</sub> receptor |
| WO 9325692     | cDNA sequences encoding human inositol monophosphate                                    |
| EPO 481673A2   | cDNA for FKBP sequences                                                                 |
| EPO 514207A2   | cDNA for human neurokinin-1 receptor                                                    |
| WO 92/18523    | Segment of DNA comprising a human cholesterol 7α-hydroxylase                            |
| EPO 496453A1   | Segment of DNA comprising a human glut 4 promoter                                       |
| EPO 419099A1   | DNA for protein inhibiting factor Xa                                                    |
| EPO 533350A1   | cDNA encoding precursor interleukin-1β                                                  |
| UK GB 2264948A | cDNA sequence encoding adenosine receptors                                              |
| UK GB 2265375A | cDNA sequence for human neurokinin-3 receptor                                           |
| UK GB 2265376A | cDNA encoding steroid hormone receptor superfamily                                      |

set a clear course for public policy. Can we now expect the company's entire collection of 'research tools', including cDNA sub-genomic sequences on which they have filed patents (see table), to be made available in public databases, including their extensive chemical-structure databanks, data on purified receptors, enzymes, cytokines, growth factors, monoclonal antibodies, immunological reagents, nucleotide probes and the myriad reagents that constitute the 'tools' of modern drug discovery? Clarification of this question, and whether an intellectually consistent position will be adopted, is awaited with great interest by other companies.

#### Future challenges

The debate about genomic patenting is likely to intensify. Opposition to genomic patents has little or no foundation under current law, other than to oppose overly broad patents that grant rights to uses neither claimed nor envisaged. The momentum of the debate comes more from forceful expression of philosophical,

theological and moral objections about the implications of genetic research and from the struggle for industrial competitiveness. Patent law is entirely unsuited to arbitrate on moral and ethical questions. A ban on genomic patents will not shield society from the evolution of genetic medicine or the need for vigilance against the misuse of genetic testing or the genetic modification of humans. Resolution of such complex issues demands more than a myopic focus on patent law and must engage the full spectrum of societal opinion. The reports on the ethics of genetic testing from the UK Nuffield Council<sup>24</sup>, the US Institute of Medicine<sup>25</sup> and the UK parliamentary Science and Technology Committee<sup>8</sup> are excellent examples of the comprehensive scholarship needed to guide protection.

The intellectual merits and economic value of patent laws, built and refined over five centuries in Europe and over two centuries in the United States, should not be dismissed without full assessment of the consequences. The potential conflicts inherent in the commercialization of genetic research, and the complexity of the social implications of genetic medicine, will still exist. We must examine these difficult topics in depth while avoiding over-simplification for popular appeal. If accepted uncritically as solutions, superficial slogans, such as 'genes are merely research tools' and 'no patents on life', will erode the foundation of intellectual property rights needed to convert innovative research into new products, to the future detriment of medicine and human welfare. □

George Poste is at SmithKline Beecham Pharmaceuticals, 709 Swedeland Road, King of Prussia, Pennsylvania 19406-5939, USA.

1. *Nature* **371**, 363 (1994).
2. *Nature* **371**, 270 (1994).
3. *Nature* **374**, 750 (1995).
4. *Nature* **374**, 391 (1995).
5. Smith G. K. & Kettelberger, D. M. *Am. Intelect. Prop. Law. Ass. Quart. J.* **22**, 28-64 (1994).
6. National Academies Advisory Group *Intellectual Property and the Academic Community* (Royal Society, London, 1995).
7. *HUGO Statement on Patenting of DNA Sequences* (Human Genome Organisation, Bethesda, 1995).
8. Science and Technology Committee *Human Genetics: The Science and Its Consequences* Vol. 1 (HMSO, London, 1995).
9. *Patenting Human Genes: A Report* (Danish Council of Ethics, Copenhagen, 1994).
10. *Human Tissue: Ethical and Legal Issues* (Nuffield Council on Bioethics, 1995).
11. *Biotechnology Products: EPLC Pharmaceutical Law Report No. 9* (Scrip, Richmond 1994).
12. Office of Science and Technology *Intellectual Property in the Public Sector Research Base* (HMSO, London, 1992).
13. Warshofsky, F. *The Patent Wars* (Wiley, New York, 1994).
14. *Realising Our Potential: A Strategy for Science, Engineering and Technology* (HMSO, London, 1993).
15. Office of Technology Assessment *Biotechnology in a Global Economy 203* (US Government Printing Office, Washington DC, 1991).
16. US public laws 96-480 and 96-517 (1980) and amendments at 15 USC 3701-14 and 35 USC 200-211 (1988).
17. Scott-Ram, N. *Nature* **364**, 666-668 (1993).
18. Cohen, J. *Science* **268**, 1715-1718 (1995).
19. *Fed. Register* **60**, 12771 (1995).
20. Butler, D. *Nature* **370**, 4 (1994).
21. Adams, M.D. *et al. Nature* **377** (suppl.) (1995).
22. Williamson, A. R. and Elliston, K. O. *Nature* **372**, 10 (1994).
23. Re Deuel *et al. U. S. Patent and Trademark Office* 94-1202 (Ser. No. 07-542, 232) (1995).
24. *Genetic Screening: Ethical Issues* (Nuffield Council on Bioethics, 1994).
25. *Assessing Genetic Risks* (eds. Andrews, L. B., Fullerton, J. E., Hotzman, N. A. & Motulsky, A. G.) (National Academy Press, Washington DC, 1994).



# China still hopeful ten years on

China's science is in much better shape than a decade ago, but there is still a long way to go. A cumbersome language and the prospect of political instability are amongst the problems to be overcome

CHINA has made great strides since *Nature's* previous survey of China just ten years ago (*Nature* 318, 205–228; 1985). But there is still much to be done before the most populous country in the world will realize its huge ambitions in science and technology. **The plan is nothing less than to treble, by the end of the present century or in just five years, the percentage from 0.5 to 1.5 per cent of gross domestic product (GDP) devoted to research and development.** If that can be done, it will be the most rapid mobilization ever of technical resources for strictly civil purposes. (Defence research in China is separate and off-bounds.)

What are the chances that China will succeed? It would be rash to jump to the conclusion that even this unprecedented goal will not be attained. What has been done in the past ten years to rebuild and extend the major cities is by itself a proof of what 1.2 billion well-disciplined people can accomplish if they have the resources to invest in buildings and industrial plant. The visual impact of change is plain for everybody to see. In Guangzhou (once called Canton), the city now almost surrounds the once-rural airport. The Pearl River flowing through it is alive with river traffic, as if it were a Canaletto painting in which T. S. Lowry had had a hand. There and elsewhere, the skyline is punctured by construction cranes, many of them putting up slab-like buildings of no merit. Hong Kong's most valuable contribution to China after 1997 may yet be its sense of architectural style.

The physical sight of Chinese cities is nevertheless not the most profound change there has been. The coming of everyday commercialism is a more obvious transfor-

mation. (Sometimes it seems to have become a rip-off society.) Ten years ago, people in the cities of the eastern seaboard of China were complaining at their impoverishment relative to the once-poor people

and guessed that the disparity would get worse before it could improve. (One of the difficulties is that the Chinese government's tax revenues, dependent mostly on what companies pay, are a declining pro-



of the agricultural hinterland; price controls had been lifted for farmers' crops, but there had been no change in the state-controlled wages paid to the likes of urban scientists and academics.

Now the boot is on the other foot. Urban centres, especially those along the eastern seaboard, are booming, while the agricultural regions (called the "rural areas" in the patronizing language of the urban populations) are sinking back into their habitual relative poverty. One result is that farmer's have taken to growing cash crops (fruit and potted plants to decorate the living rooms of the urban population, for example), putting China's self-sufficiency in cereals (especially wheat) under strain for the first time in a quarter of a century. And the urban railway stations are thronged by newly arrived migrants from the rural areas, seeking a share of the action in the cities.

**The economic gap between the cities and the rural areas is getting wider all the time.** A comment by the World Bank recently estimated that the urban populations of China are, on the average, twice as prosperous as those still in the rural areas,

portion of GDP.) It seems generally agreed that the growing disparity presents the government with a serious political problem. Totalitarian the government may be, but there are limits to the degree to which the impoverishment of the estimated 700 million people still dependent on agriculture can be perpetuated without an explosion of some kind.

The government is seized of the problem, and has a strategy for its solution. Grade-school education (up to 16 years) is one priority, remedying persisting adult illiteracy is another. But those caught in the rural poverty trap who most need help from services like these are prevented from enjoying them by the circumstance that educational services are not free of charge. A further impediment is the sheer difficulty of learning the Chinese ideographs, which necessarily (as in Japan) occupies several years of early schooling, raising the question whether China can successfully modernize itself without devising a more manageable script.

That is an issue for those writing software for the growing computer businesses. The ideographs are broadly the same, with

## Science in China

This survey has been written by John Maddox, Editor of *Nature*, and David Swinbanks, Tokyo Correspondent and Publisher of *Nature* in Japan. It is based on visits to Shanghai and Beijing in September (D.S.), Guangzhou, Shanghai and Beijing in October and November (J.M.) and to Nanjing and Beijing in November (D.S.). *Nature* is grateful to the Academy of Sciences for arranging most of the visits as well as for transport (which was paid for) as well as to numerous academics and researchers for their time, trouble, hospitality and good fellowship.

the same meanings, as those used in Japan but there are many more of them. As with Japanese, when phonetic representations of Chinese words are entered into a word-processing program, the user is required to choose its correct representation from a string of a dozen or so ideographs displayed on the screen. The process is slow, prone to error and does not obviate the need for all those years of schooling.

There is no real likelihood that China will abandon Chinese for some other language (English for example); the cultural tradition is simply too vivid and too strong. But the view often stated by Chinese scientists that it is only a matter of time before colleagues elsewhere will be learning Chinese so as to find out what appears in Chinese journals misses the point about the language: when the content of written Chinese is related only by elaborate conventions to the spoken language, the problems of communication stem not from those of translation but from the investment of time and memory required to write and then to read. That will threaten the goal of universal adult literacy for a long time to come.

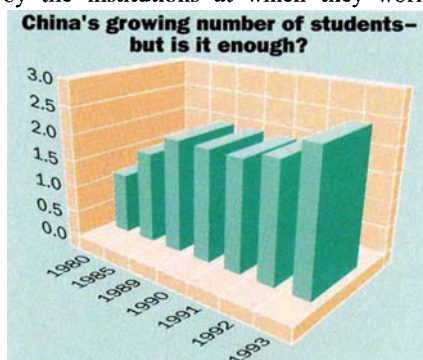
In China's education policy, literacy (including grade-school education) takes pride of place for good reasons, many of them political. By comparison, the universities are neglected still. **Although student numbers have been growing steadily, the total student population now amounts only to 2.5 million. That compares with 1.8 million in the United Kingdom and 2.2 million in France, both of them countries roughly 5 per cent as populous as China. In short, China has many fewer students at its universities than comparisons with industrialized countries elsewhere would suggest appropriate.** Is that a problem?

There are conflicting views. One senior academic holds that the comparison with countries elsewhere is misleading, because people from "the rural areas are not interested in a university education". Others argue that skills of the kind that China most urgently needs are technical skills, best provided by on-the-job training at production plants and research laboratories. But others acknowledge that even the major universities in China, reformed though they have been in the past ten years, are less flexible than modern circumstances require.

Inflexibility arises in several ways. One is the general convention that universities must provide housing both for faculty members and for students. Both are acknowledged as necessary. Students' families usually have too little living space to allow for private study, while daily travel would be impossible, certainly for bright people from the rural areas. The troubles thus caused are not merely the extra capital cost of housing, and the need for land, but the managerial demands of catering for students' living needs.

China may be a communist state, but university education is not free. Costs vary from one university to another, but board and lodging may exceed 2,500 yuan a year, with a further 1,000 yuan for tuition. These are not large sums in absolute terms (roughly US\$400 a year in aggregate), but they are still large sums for most Chinese families, particularly those in rural areas. The Chinese zeal for education notwithstanding, they are probably an impediment to any proposal for the rapid increase of the student population.

Graduate students, whether at universities or research institutes, fare better. They are almost invariably paid a small stipend by the institutions at which they work,



much in the spirit that graduate students in the United States and elsewhere may sometimes be paid as research assistants and thus be charges on research grants held by senior people.

Faculty housing, once necessary because of the physical shortage of housing in the major cities, is now a means of supplementing teachers' salaries (routinely between 600 and 1,000 yuan a month), rather less than the monthly rent of the apartments springing up everywhere in the major cities. Faculty members pay only nominal rents to the institutions that employ them. The difficulties here are not only administrative, but cause faculty members to become trapped in the universities to which they are first appointed.

These impediments to change have fortunately been largely eroded by the economic free-for-all in which institutions of all kinds now make their way in the new China. Universities and research institutes, free as they have become to set up joint ventures with commercial interests, have ways of shunting members of staff who may have become intellectually sleepy into employment that is less directly a brake on what departments are attempting to do. (Those asked to take on this work have a right to return if the enterprise gets into trouble; those who elect to go for larger salaries must cut the cord.)

**Commercial ventures have given many of the major universities substantial untied income and a measure of autonomy as a consequence.** The idea that the University of Peking (also called Beijing) should derive roughly a quarter of its

income from the profits of a company that markets electronic typesetting software and equipment to Chinese newspaper printers will startle many academic institutions elsewhere, but the pages that follow will demonstrate that this is not a rare exception.

The emergence of the business-based university fits in with the government's exhortation to the academic community that it has a responsibility to be "socially useful". So far, the management of industrial enterprises from within university departments seems not seriously to have damaged the quality and the character of academic life.

But these are early days, when the physical economy is so short of modern products and services that most ventures have a high chance of success. The Chinese economy is not merely growing quickly (at 10–12 per cent a year in real terms), but in qualitative terms it is being created from scratch, on a virtually blank slate. It is difficult to foresee what will happen when academically owned venture companies encounter effective competitors created within the newly emerging private sector. And while the first wave of business-based universities is proud to boast of its achievements in the commercial field, it is not easy to see how the same recipe could be followed by many of the other 1,000 universities in China outside the select company of the 36 directly responsible to the State Commission on Education.

The ability to play the economic markets has given the successful Chinese universities a degree of *de facto* autonomy they lacked ten years ago. What state commission can argue with success? Yet formally, matters such as the appointment of university presidents remain in the hands of the State Commission. That is why it was of more than passing interest that one of the smaller universities in Shanghai took the daring step, at the end of October, in electing its own president before sending the name to the State Commission for approval. Self-confidence of that kind seems destined to flourish.

**Another crucial change to have transformed the Chinese university system is that the once-established gerontocracy has been dismantled.** To some extent, this has been forced on China by the consequences of the Cultural Revolution, when, for the best part of 15 years, many academics (and especially those with previous connections in the West) were compelled to carry out high-level community service of one kind or another. The result is a gap in the age-spectrum of academics in post of those older than, say, 45. That provides the opportunity for appointing to senior posts, as heads of department or in university administration, young men and women who have acquired an academic reputation elsewhere. Those with a research background in North America or in Europe are



much sought after.

It cannot be long before these young turks begin to affect the working of the university system, and of the research enterprise as a whole. Their views of how science should be practised are formed in the West, and are largely Western. They publish in international journals by preference, they compete among themselves for graduate students and they are skilled at accumulating research funds. There are enough of them to make it likely that, ten years from now, the 36 key universities will be institutions of international class, but still not cosmopolitan universities in the strict sense. **Sooner rather than later, China will need mechanisms for attracting scholars other than ethnic Chinese to its academic institutions.**

Meanwhile, the great Chinese diaspora, stretching from Singapore to Los Angeles and New York, is a great source of strength. The liberalization of the "socialist market" seems to have caught the imagination of ethnic Chinese throughout the world. They are ready to entertain proposals for research collaboration and joint appointments. As one such (working in Singapore) said, "China now is like the old Wild West in the United States, full of opportunities and risks".

Occidental collaborations are nevertheless far from negligible in their importance. The Chinese Academy of Medical Sciences is still housed in the Union Medical College in Peking, founded and built by the Rockefeller Foundation in the 1920s. And everywhere there are well-established academics who were Humboldt Fellows in Germany in the 1960s and 1970s, and whose links with Germany remain productive and robust.

But links with the United States remain the strongest. The thousands of Chinese who took advantage of the huge scholarship schemes of the 1970s, and who then stayed in the United States, are potentially a huge resource. Even if they stay where they are, academics at Chinese institutions will look to them for continuing collaboration and other assistance. That is why there seems to be comparatively little concern in modern China about the high proportions of those sent to the United States on PhD courses or as postdoctoral fellows who fail to return. (The attrition is especially marked in fields such as computer science, organic chemistry and molecular genetics.) Once Chinese, always Chinese, seems to be the motto.

**The improvement of research at the major universities in the past ten years has righted the old imbalance between academic institutions proper and the government research institutes.** Although the great falling-out between China and the former Soviet Union took place in 1958 or thereabouts, the organization of science in China was, at the outset, modelled on the Soviet system. The principle was that universities

would train undergraduates, that research institutes would be responsible for the bulk of research (and even much industrial development) and would be the chief means of training young people in research.

Even ten years ago, the Chinese Academy of Science was the most influential voice in the organization of research, although its research institutes (now 123 in total) were never as huge as those of the Soviet system. Now the academy's institutes are required, like the universities, to fend for themselves except for the provision of a basic budget. The competitive grants schemes that have given university departments opportunities to succeed are also open to applications from the research institutes, as is the opportunity to found joint ventures with commercially interested partners.

The system, unique to China, whereby laboratories in universities and research institutes are designated as national 'key' laboratories (and blessed with extra funds as a consequence) is potentially of great importance. It is probably too soon for those concerned to have considered whether the end result may be that some key laboratories overshadow the institutions from which they sprang.

Many (but not all) of the academy's research institutes have as a consequence become more like the sources of industrial innovation that their founders foresaw nearly half a century ago. Yet, like the better university departments, their most important role may still be that of training graduate students, which are now selected on the basis of results in a national competitive examination. (The old system by which all university graduates were directed to their first jobs by government agencies has been swept away, to the general contentment, together with the old traditions that party cells would meet on laboratory premises on Saturday mornings to discuss matters of political interest to the government.)

Research programmes naturally vary from one place to another, but it is still the case that much of the work at the research institutes consists essentially of the replication of what has been published in the open literature elsewhere. That, of course, is a way of ensuring that China remains in touch. The test of success, ten years from now, will be whether the number of laboratories at which research breaks new ground can still be counted on the fingers of a few hands. As things are now, Chinese researchers are confident that they will be up there with the leaders. Visitors, more aware of wider issues, may think it will take a little longer.

**The most serious uncertainties lie not within the scientific enterprise, but are political.** To listen to invariably optimistic Chinese, it is not easily divined that China was at war for nearly 30 years from the late

1920s with invaders (from Japan), that the war was ended in 1948 only after a civil war (and the Long March), that the government since then has survived several traumas (the Cultural Revolution and the killing of more than 200 dissidents in Tiananmen Square) and that the current interest is what will happen when Deng Xiaoping, the leader without formal title, eventually dies. Deng has been in poor health for some time and has not been seen in public for more than a year.

A classic power struggle for the succession to Deng is less likely than a bout of recrimination about the reasons for the repression in Tiananmen Square. Nobody acknowledges knowing what happened to the student who defied a tank during the confrontation, but there is a general unease about the whole incident that will have to be stilled if the present pragmatic government is to continue on its present pragmatic way. And that process could be traumatic.

China's external relations are also potentially a source of instability. The status of Taiwan is an immediate issue. Left to itself, the government of China in its present mood would probably prefer to see how the *anschluss* with Hong Kong in 1997 passes off before raising the question of how Taiwan should be administered. (In China's eyes, what it professes to regard as another province should in due course be administered from Beijing, as Hong Kong will be in less than two years.) Economic success in China would also help to tempt Taiwanese to let their allegiance to the diaspora overcome their understandable reluctance to accept China's view of their status. But anything could go dangerously wrong with that calculation, changing many well-laid plans.

Hong Kong itself is not an issue, at least to working Chinese. Hitherto, except for academics and government officials, hardly anybody from the mainland has been there except on a one-way ticket. Hong Kong's belief that its riches make it a kind of endowment to the mainland, and that its overseas connections make it a bridge to the West, is not much echoed on the mainland. Nobody should be surprised if the government of China dutifully sticks to the letter of the Basic Law regulating the transfer, with its 50-year moratorium on institutional change, while seeking to make changes in Hong Kong's way of life by persuasion.

Internal trouble seems less of a potential threat to stability. At least along the booming east coast, people are content with the signs of change, while Guangdong (the province of which Guangzhou is the capital) was chafing at the central government's uncommercial ways as recently as ten years ago. The seriousness of the tension supposedly building up between the rural areas and the cities could not have been encountered in journeys between eastern seaboard cities. □

# Competition and ventures all the rage

THERE have been dramatic changes in the way government research is funded and managed in China. Ten years ago, at the time of *Nature's* previous survey of science in China, the first organization for a nationwide system of competitive research grants, the National Natural Science Foundation of China (NSFC), was just in the process of formation. Now there are many such programmes funded by various branches of the government and open to a broad spectrum of researchers, in both universities and research institutes of the various government-run academies, the Chinese Academy of Sciences for example.

With inspiring names such as Gongguan, meaning 'storm the strategic position', Pandeng, or 'climbing (a mountain)' and Xinghuo, the 'star fire' or 'spark' programme, or simply named after the time of their establishment, such as the '863' programme initiated in March 1986, these

grant schemes have created a new competitive environment in which some research groups are now "rather rich", says Xu Yonghua, deputy director of the academy of sciences Institute of Cell Biology in Shanghai.

The multiple sources of funding have given greater autonomy to universities and research institutes. Now they have to "compete in the market to survive", says Lin Guo-Qiang, director of the academy's Institute of Organic Chemistry, also in Shanghai. This is a complete turnaround from the era before the NSFC was formed, when each government ministry and agency fed its own research organizations with small amounts of money, and large projects were assigned in a 'top down' manner.

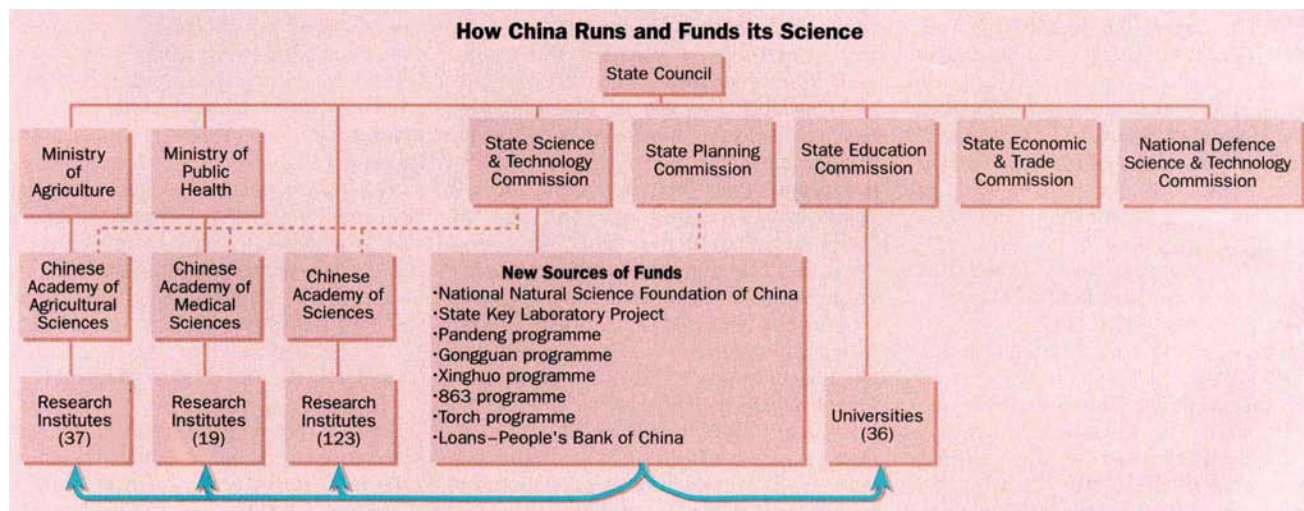
Previously, China had a "Soviet style" system with research concentrated in large government institutes, such as those of the Academy of Sciences, and universities

were "very poor", says Feng Duan, professor of physics at Nanjing University. But now the universities can compete on equal terms and they are doing "a lot of good research", often "much more efficiently" than the government institutes, Feng claims.

The institutes of the academy of sciences are feeling the heat. "Ten years ago, we got limited but guaranteed funding", says Nan Rendong of the Beijing Astronomical Observatory. Now he worries "day and night" where the next budget is going to come from.

Xu at the Institute of Cell Biology, who is a cancer researcher, got his first grant of 50,000 yuan (US\$6,000) from NSFC in 1984, just after he returned to the institute from a two-year stay at the National Institutes of Health in the United States.

Before 1984, all funds for his institute, amounting to less than 2 million yuan a year, came from the academy. But now



The State Science and Technology Commission is the key organization for formulating and implementing science and technology policy within guidelines set by the State Council, the highest body for day-to-day running of the country headed by Premier Li Peng. An exception, however, is the National Defence Science and Technology Commission which runs its own independent military research. The commission is responsible for managing the many new competitive funding programmes set up over the past decade, including the National Natural Science Foundation, the State Key Laboratory Project, the "863" programme, the 'storm the strategic position' or Gongguan programme, the 'climbing' (Pandeng) programme, the 'spark' programme for agricultural research, and the 'torch' programme for developing high-technology zones. The commission also selects research organizations for loans from the People's Bank of China under reforms of the bank introduced in 1985. In addition, the commission supervises the flow of operating funds to the various academy research institutes, and research institutes of other ministries (not shown in chart) and it is responsible for coordinating research that spans these various research institutes. This relationship is indicated by a dashed line in the chart. But the Chinese Academy of Sciences enjoys a considerable degree of autonomy. For example, it can win approval direct from the State Council for funds to import expensive scientific equipment, and the science and technology commission cannot change the fixed oper-

ating funds allocated to the academy by the ministry of finance. Similarly the Academy of Medical Sciences and the Academy of Agricultural Sciences, have a fair degree of autonomy from the commission and are overseen by their respective ministries. China's universities also enjoy a considerable degree of autonomy. Thirty six of the more than one thousand universities in China, including the most prestigious ones, such as Beijing (Peking), Qinghua, Nanjing and Fudan universities, come under the jurisdiction of the State Education Commission, which was elevated from ministry to commission status in the reforms a decade ago. There are many other universities run by various ministries (not shown here), and the Academy of Sciences has its own University of Science and Technology. The universities receive operating funds from the Education Commission or their respective ministry (or academy), but they can also obtain considerable support from the various competitive schemes run by the Science and Technology Commission, and from the private sector. The State Planning Commission, which is responsible for managing the national economy, also plays an important role by funding infrastructure for major science and technology projects, such as the buildings of state key laboratories, high technology zones, and pilot plants for industrial scale-up of new technologies. Similarly, the State Economic and Trade Commission supports programmes to modernize industrial technology and oversees the import and transfer of technology.

the institute spends about 40 million yuan a year, only 2.8 million yuan of which comes as operating funds from the academy. The rest is derived from competitive grants, contracts with industry and from venture companies set up by the institute. Such venture businesses are another important source of funds in the new competitive era.

The situation is similar at many of the academy's 123 institutes and at China's leading universities as well. Ten years ago, the academy's Institute of Remote Sensing in Beijing, for example, got 72 per cent of its 2.2 million yuan budget from the academy, but last year only a little over a third of its 19.2 million yuan budget came from that source.

At Fudan University in Shanghai, only about 10 per cent of the 30–40 million yuan annual budget now comes from the State Education Commission. A third to a half comes from contracts from private enterprises — for example, a computer control system for a power plant in Shanghai — and a similar amount from the new government programmes.

At present, the general grants of the NSFC, which support basic research by thousands of researchers in universities and research institutes, are typically worth 80,000–100,000 yuan (US\$10,000–\$12,000) each per year. There are also some "key" grants for 0.5 million yuan and a few "top" grants of 1 million yuan per year. All in all the foundation awarded 584 million yuan in 1994.

Another characteristic of the past ten years has been a swing to more applied research. While the foundation primarily supports basic research, most of the new grant schemes target more applied research with the aim of boosting growth of the market economy or advancing agriculture and medicine.

In March 1986, a group of scientists wrote to Deng Xiaoping suggesting the set up of what subsequently became the '863' programme to support research and development in the fields of space technology, information technology, lasers, automation, energy, new materials and biotechnology, in particular medical and agricultural applications of biotechnology and protein engineering. Run by the State Science and Technology Commission, the programme awards grants for 5 years that range up to several million yuan in value. They are more difficult to get than those of the NSFC and are "very competitive", with a less than 20 per cent success rate for applications, says Xu (who has received one).

The commission has also set up the Gongguan ('storm the strategic position') programme to fund applied research with grants in excess of 1 million yuan per year, and the Pandeng ('climbing') programme to focus on research of a more basic yet still applied character — strategic

research as it might be called elsewhere.

The competitive environment has been stimulated in a particularly influential way by a system set up by the State Planning Commission in 1984 for targeting large sums of money towards key laboratories at universities and institutes. The idea is that institutions should vie with each other for designation. The scheme is open not just to universities but to institutes of the academy of sciences, and the academies of medical sciences and agricultural sciences, belonging, respectively, to the ministries of health and agriculture.

About 150 state key laboratories are now supported under the scheme, around a hundred of them in universities. Several of those of a more practical and applied character, such as the State Key Laboratory of Severe Storm Research at the University of Peking, are supported by a World Bank loan.

Each key laboratory gets a large budget at the start of its operations. For example, the State Key Laboratory of Surface Physics at the Academy of Sciences Institute of Physics in Beijing had an initial budget of US\$3 million plus 3 million yuan to buy equipment and facilities; the continuing annual budget is 1 million yuan.

Every three years, the key laboratories undergo an external evaluation organized by the NSFC, and laboratories are ranked according to these evaluations. Additional funding is allocated to those judged excellent and key laboratory status is withdrawn if a laboratory fails to meet standards in two consecutive evaluations.

Laboratory directors and the chairmen of the independent committees that supervise each laboratory are appointed on limited terms, and laboratories are encouraged to employ scientists from overseas as well as from China.

Such practices have spread beyond the state key laboratories. Some institutes in the academy of sciences, for example, now employ researchers on short-term contracts, and staff undergo regular assessments every few years. This has led to some research groups being closed down or reduced in size.

At the academy's Institute of Physics, for example, the staff complement has been reduced from 1,200 to 700 in ten years. Only about 200 people of the 500 who have been shed have retired. But during the same period, the same institute has also established many new groups for research in new fields such as superconductors, surface linear optics, nanostructures and femtosecond phenomena, and there has been a surge in the number of graduate students from fewer than 100 ten years ago to 160 at present.

The systems of assessment are quite rigorous and complex. They are constantly modified and refined and vary from organization to organization. But all use the

number and quality of publications as a measure of success.

At the Beijing Astronomical Observatory, for example, publications are divided into numerous categories. For example, a paper in a leading international journal such as *Nature* or the *Astrophysical Journal* wins 400 points for the author(s). Papers in run-of-the-mill Chinese journals rate only 50 points, but a paper in the *Chinese Journal of Astrophysics*, which is quite well respected, earns 100 points. Journals listed in the *Science Citation Index* get at least 250.

Invited oral presentations win points depending on the fame and importance of the meeting. Points are also allocated for services, such as organizing an international conference, or chairing a committee or society.

At the end of this process, points are added up over a 3-year period, and the score is recalculated every year. The outcome is used to determine a monthly bonus, paid on top of the low monthly salary (typically between 600 and 1,000 yuan). Sometimes, the bonus can exceed the standard monthly wage. Those who emerge with poor scores, on the other hand, find that life becomes very hard. Such people are encouraged to go elsewhere — often into venture companies linked to the observatory.

Overall, leading universities in China have benefited most from the shift to competitive grant support and the new-found plurality of sources. Ten years ago, the universities were poor cousins of the institutes of the Chinese Academy of Sciences, depending as they did on the rather conservative State Education Commission for research support.

Now, as well as earning a substantial grant income, many of the leading universities have thriving off-campus businesses that bring in substantial amounts of extra funds. The University of Beijing and the Qinghua University (both in the capital, Beijing) are major participants in a new high-technology zone established in the city on land owned by the Academy of Sciences. The academy claims the zone has more than 400 venture companies, although officials admit that only a few tens of these companies are of any substance. One of them is nevertheless the computer typesetting company that sells hardware and software to Chinese printing plants and from which the University of Peking derives a large slice of its total budget.

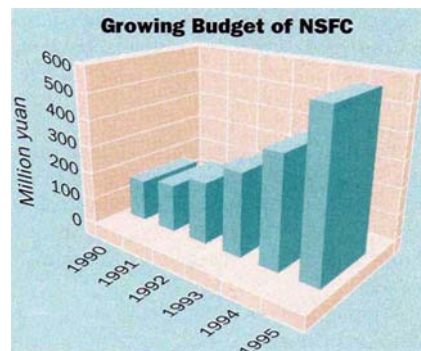
On the same principles, the Fudan University in Shanghai set up Fudan Development Company of Science and Technology in 1985 in response to the government's urging that universities should "serve society". The company later changed its name to Fuhua and it now employs about 300 local people, including 40 members of the university faculty. It



## China sets ambitious goals for R&D

Earlier this year, China set ambitious (some would say impossible) goals for boosting its investment in research and development (R&D).

In 1994, China spent only 0.5% of its gross domestic product (GDP), or 22.2 bil-



lion yuan (US\$2.7 billion) on R & D. (This excludes expenditure on military research, which is not made public). By comparison, nearby Japan invests about 2.8% of its GDP. Furthermore, the percentage for China has declined from 0.7% in 1992, despite rising outlays for R&D, because of the rapid growth of the Chinese economy.

In May, the State Council headed by Premier Li Peng announced plans to reverse this trend and increase R&D expenditure to 1.5% of GDP by the turn of the century, only five years away. Industry is expected to play a crucial role in the spending drive.

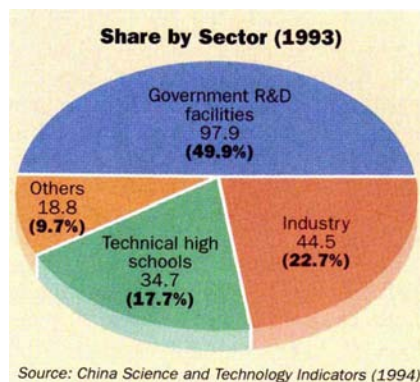
In 1993, only 23% of outlay was from industry (see pie chart). But the government is aiming to increase this share to 50% by 2000, says Li Maoming, deputy secretary general of the State Science and Technology Commission. She explains that this will be done with tax incentives.

For example, if a company increases its

research and development expenditure by 10% it will get a 150% tax reduction. Furthermore, for technology transfer valued at less than 300,000 yuan (US\$36,000) companies will no longer need to pay tax on the income they earn from such transfers.

Li is confident industry will play its role. In the planned economy of the past, she says, the government sold products for industry and companies were indifferent to R&D, or, as Li put it, "the daughter of the emperor does not worry about getting married". But in the present socialist market economy, there is "competition" and companies "need" to do R&D.

The government also plans to make dramatic increases in its spending. By 2000, the government intends that 10% of R&D outlay should go to basic research. One important organization supporting such research is the National Natural Science Foundation of China (NSFC). The foundation's budget has been growing quite rapidly, and now stands at 584 million yen (see graph). But even more rapid growth will be required if the foundation is to achieve its goal of providing about one third of all outlays for basic research by 2000, or more than 2 billion yuan.



makes unstopped power sources (UPSs) for computers, batteries, intelligent cards and biotechnology products.

Last year, the company made sales of 200 million yuan (US\$24 million) and profits of 21 million yuan. With a shareholding of more than 30 per cent, the university's dividend amounted to 8.1 million yuan, which it ploughed back into the improvement of buildings and other infrastructure.

Most of the research institutes of the Academy of Sciences also run their own businesses. Beijing Astronomical Observatory, for example, has set up the Asia Satellite Company to sell satellite receivers for televisions all over the world. With an annual turnover of 15 million yuan and profits of 2-3 million yuan, the company returns about 0.5 million yuan to the observatory each year.

Close to 100 observatory staff out

of 358 now work at this company and eight others set up by the observatory. The observatory provides only housing and welfare for these employees; otherwise, company staff are entirely self-financed. By 2000, the number of such employees will have further increased; meanwhile the observatory hopes that it will have shrunk its core staff to about 100 people.

The downside of these changes is that there has been a clear swing towards applied research. This is the aim of government policy which links science and technology to development of the 'market economy' and demands social benefits from research. At Fudan University, for example, all the newly established departments in the past ten years have been on the applied side; there have been new departments in computer science, microelectronics, materials science and genetic engineering, for example.

Departments involved in more basic research are struggling to survive. For instance, the department of nuclear physics now has no undergraduate students, only graduates. But Zhang Ju Xiu, deputy director of the university's department of science and technology, claims everyone is agreed on the need for reform and that reform "must be based on the needs of society; so we are expanding applied research and just keeping small teams of basic researchers".

The University of Nanjing appears to have been more successful in maintaining a strong component of basic research in the changing climate. Its physics department is ranked first in China and it has the biggest and oldest department of astronomy. The university also has seven state key laboratories, many of them doing quite basic research, such the Laboratory of Solid State Microstructures (see page 543).

The university's president, Qin-yue Qu, says Nanjing has kept "two hands ready" by developing applied research of immediate value to society and then ploughing some of the profits back into basic research.

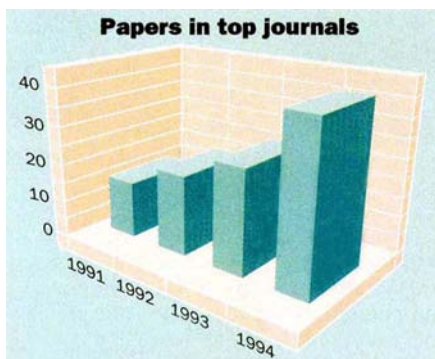
Young researchers have benefited greatly from the new environment. Government policy now calls for the cultivation of young researchers and both universities and government institutes are offering senior positions, such as directorships of institutes and professorships, to young Chinese scientists still in their thirties. This is a complete reversal of the situation a decade or so ago, when Confucian tradition dictated that the top positions should be held by senior scientists in their 50s, 60s or 70s.

The result is that many bright young Chinese scientists are being attracted back from the United States and Europe. One example is Fangzhen Sun, who has recently been appointed director of the academy's Institute of Developmental Biology, after 8 years at the University of Cambridge in Britain. Another is Mu Wang of the National Laboratory of Solid State Microstructures at the University of Nanjing, who is (in his early thirties) one of the youngest professors in the university. Wang came to the laboratory from the University of Nijmegen in the Netherlands.

Both Wang and Sun hold 'premier' grants, a new and very competitive grant scheme for young researchers under 45 given in the name of the premier Li Peng. Wang's award is worth 0.6 million yuan and he says that about 500 young scientists compete for about 50 of the awards each year. This influx of young scientists, many of them from the West, may do more than anything else to stimulate the rapid development of China's science. □

## Centre of excellence on the Yangtse

**Nanjing** Feng Duan, one of China's leading physicists, spent part of the Cultural Revolution laying asphalt on a huge double-decker bridge that spans the Yangtse River at Nanjing. Feng says he felt



Papers in *Nature*, *Physical Review Letters*, *Applied Physics Letters* and *Journal of Applied Physics* by authors from the Laboratory of Solid State Microstructures.

a great sense of achievement when he stood on top of the completed bridge with his comrades. Nearly two decades later, he seems almost as proud of the bridge as he is of his Laboratory of Solid State Microstructures at the University of Nanjing, which is ranked as one of China's top key state laboratories in the field of physics.

The laboratory, set up in 1984 as one of China's first key state laboratories, has been producing increasing numbers of papers in leading science journals in recent years (see graph). All 34 state key laboratories covering physical and chemical sciences have just undergone an external evaluation organized by the National Natural Science Foundation of China. Although the official results are not yet out, the Nanjing laboratory is now ranked among the top three, and ahead of the much bigger Institute of Physics at Beijing belonging to the Chinese Academy of Sciences long considered China's leading physics research institute.

The Nanjing laboratory's research ranges from simple but elegant studies of crystal growth using a standard light microscope and a video recorder (see, for example, *Nature* **367**, 438; 1994) to sophisticated atomic-scale nuclear magnetic resonance studies of spin polarization in interlayered magnetic and non-magnetic materials (*Phys. Rev. Lett.* **72**, 768-771; 1994). The laboratory also claims to be the first in China to have produced quantum wires and wells.

Feng's group is unusual in that experimentalists and theoreticians work together. Traditionally, in China (and elsewhere), theoreticians tend to work on their own.

Many young physicists have been attracted back from overseas to work in the laboratory, which has more than 30 perma-

nent research scientists and even more visiting scientists. They like the dynamic interaction between physicists of different backgrounds, the strong financial support from the government and the fact that there are several leading scientists like Feng in the laboratory.

One such researcher is Mu Wang, who is working on crystal growth. He set up a laboratory with optical microscopes and a video system with a grant of 150,000 yuan for 3 years under the 'climbing' (Pandeng) programme before going to Nijmegen University in the Netherlands in 1992 on the recommendation of his professor, Nai-ben Ming, who has now succeeded Feng as director of the laboratory.

On his return last year, Wang became, at the age of 31, one of the youngest professors in the university. He is now receiving even more support of 890,000 yuan (more than US\$100,000) over three years through various grant programmes, in particular a

'premier' award for young scientists. He also now has several students working under him.

"It would be impossible for me to get an equivalent position in Europe at my age", says Wang in explaining his return to China. "Besides, the academic environment in this laboratory is good and there are many active interactions between researchers working in different fields of physics, chemistry and materials science."

Wang's unusual (and only) complaint is that he does feel pressure to link his research to potential applications, in line with the government's policy of tying research to development of the market economy. But, despite the fundamental nature of much of the research at the laboratory, there are many obvious applications. And Wang may not have long to wait. Last month, a representative of Mitsubishi Materials Corporation of Japan was at the laboratory, seeking possible collaboration on the recommendation of several professors at Tokyo and Tohoku universities in Japan. □

## Nanjing University: Striking the balance

THE University of Nanjing in the ancient capital of Nanjing is a model among China's universities of how basic research can still thrive in the new market economy.

In the early 1980s, the university did not even rank among the top ten universities in China. "We were very disappointed at the time", says Feng Duan, of the university's physics department. Now Nanjing can count itself among the leading three or four universities, alongside Beijing and Qinghua, in large part because of its success in basic research.

The university boasts seven state key laboratories. Among them, the Laboratory of Solid State Microstructures (see above) and the Laboratory of Coordination Chemistry rank among the best in the country. The Department of Astronomy is the oldest, largest and strongest and has won substantial support under the Pandeng or 'climbing' programme of the central government. The School of Geoscience is also one of the best in China. And, since 1992, the university has published more papers in journals listed in the *Science Citation Index (SCI)* than any other university in China.

Feng says that much of the university's success can be credited to its president, Qinyue Qu, an astronomer. Qu has led the university for the past 12 years, an unusually long period. Normally, university presidents in China retire after a maximum of two 4-year terms. But Qu was voted in for a third term "because he was doing such a good job" says Feng.

Qu has managed to strike a healthy balance between basic and applied research. The university runs several successful companies, including a factory that produces raw materials for pharmaceuticals, and several joint ventures with foreign companies. The university's research budget has grown from 30 million to 50 million yuan over the past two years, which is ahead of the inflation rate of approximately 20 per cent; about a fifth of these funds comes from industry.

But Nanjing differs from most other universities because it has a policy that



The administrative building, Nanjing University.

some of the money from industry is given as bonuses of few thousand yuan to professors doing good basic research, says Feng. Another key to success has been a concerted effort to recruit young scholars.

But perhaps the most important driving force has been a burning desire to move up the ranks of China's universities. "Beijing University is number one, no-one disputes that, although recently we have been publishing more *SCI* papers", says Feng. "But they have no incentive to do better. We have." □



## Research institute on company lines

**Shanghai.** Lin Guo-Qiang runs the Shanghai Institute of Organic Chemistry of the Chinese Academy of Sciences as if it were a private company. That is not so surprising. Lin worked for what is now SmithKline Beecham Laboratories in the United States before becoming director of the institute.

Lin is setting precedents for the academy with his private-sector approach to management, and it seems to be paying off: the institute has made several technology transfer deals with large foreign companies; its buildings, although shoddy and badly in need of a new lick of paint, boast several gleaming new nuclear magnetic resonance (NMR) machines that together cost millions of dollars.

The institute also has a key state laboratory that receives substantial extra funds from the central government. And there is a quickly growing army of nearly 200 graduate students whose quality and number are attracting joint collaboration with researchers at Singapore National University.

All scientists joining Lin's institute have to sign a confidentiality clause and countersign it on their departure. The objective is to try to ensure that no commercially valuable knowhow leaves the facility. "I got that idea from SmithKline", says Lin. All of the 950 staff, including Lin, carry ID cards with a magnetic strip that they run through a computerized clocking-in and clocking-out system whenever they enter or leave the facility. That way Lin and the institute can keep tabs on everyone.

Under a system introduced in the early 1990s, all new staff are hired on 3-year contracts and undergo rigorous review before a decision is made on whether to extend their employment. "Almost every other week, I sign a release form for leavers", says Lin, who adds that he tries to keep the good ones by offering them a house. This is a far cry from a decade ago, when all academy institutes had permanent staff and no-one could be asked to leave.

Lin's approach does not seem to deter recruits. There is strong competition for new positions and the number of graduate students has doubled since 1990. Growing numbers of staff have recently started to join from overseas, in particular the United States and Europe, and Lin hopes to recruit 25 more young people from overseas by the end of this decade.

There have been several technology transfer and joint venture deals with large foreign companies. For example, the institute has a cross-licensing agreement with a US company (unnamed) for a beta form of polypropylene developed at the institute that is stronger and easier to mould

than the conventional form of the plastic. A joint venture with Volkswagen, one of the most successful foreign car manufacturers in China, is making anti-freeze following an institute recipe. And the US car manufacturer Ford has licensed a non-stick coating that the institute developed.

The institute also runs its own self-financed pilot plant with 450 employees for scaling up the production of chemicals. Lin sits on the board of this factory as well as running the institute.

The agreements with foreign companies are clearly bringing in a lot of extra cash. The institute's NMR machines range in price from about US\$200,000 for the smallest to US\$2.5 million for the prize machine, a superconducting one housed in its own special building.

Hsieh An Kong of the department of chemistry at the Singapore National University is a regular visitor to the institute, where he holds a joint appointment. He says the key attraction is the enormous amount of manpower available at the Shanghai institute. Few (if any) organic chemistry departments or institutes elsewhere in the world can claim such a large number of high-quality graduate students, he says.



Lin Guo-Qiang in front of Shanghai Institute of Organic Chemistry.

Nevertheless, much could be done to improve the institute. The walls of most of the buildings are peeling and cracked and the floors are dusty — hardly an ideal environment for research of any kind. The library lacks some important journals, chemicals are vented out of vertical shafts running up the sides of laboratory buildings, and the smell of organic chemicals pervades parts of the campus. Rather than more NMRs, the institute could do well to invest in some new buildings and other basic facilities. □

## Remote sensing applications

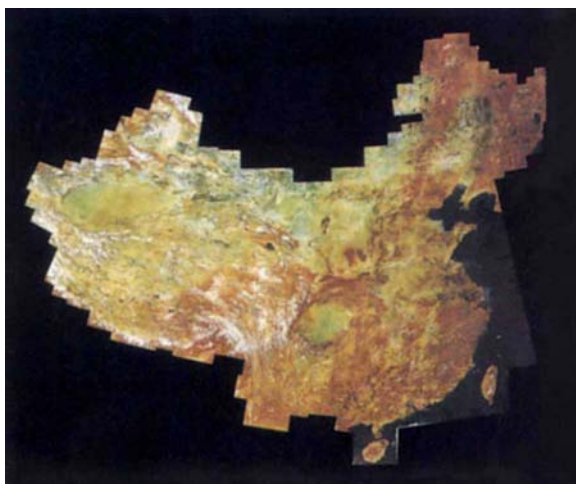
**Beijing.** In the busy eastern coastal cities of China it is easy to forget that most of the country lies west in a vast remote region stretching from the Himalayas in the south to the Gobi Desert in the north. The com-

For example, last summer, when the worst flooding in a century occurred in Jiang Xi province in north-east China, the institute used its own synthetic aperture radar (SAR) technology and aircraft to map the flooded region in 2 to 3 days. This was then applied in evacuation plans, says Liu Ji-Yuan, deputy director of the institute.

The institute collects NOAA satellite data daily with a receiving station on its top floor. The satellite images are, for example, used for drought analysis in the agricultural region of the North China Plain which fringes a desert. "More and more government officials know about our institute. And before they make decisions, they consult us", claims Liu proudly.

The institute's work also has important applications in mineral and oil exploration. For example, the institute has worked with the Texaco oil company to pinpoint oil microseepage in the Tarim Basin in the far remote western region of China. Nearer home, its remote sensing data are also used in coastal management, for example to monitor the rapid growth of the Yangtze delta off Shanghai.

Given all these applications, it is not sur-



Big target: remote sensing is one way of reducing China to size.

paratively new and small Institute of Remote Sensing Application in Beijing nevertheless has to keep an eye on it all on a day-to-day basis.

The institute, which was set up by the Chinese Academy of Sciences in 1980 and which has 254 staff, has some recent successes to its credit, and local governments in China are increasingly turning to it for advice and help.



prising that more than 60 per cent of the institute's annual budget of about 20 million yuan (US\$2.4 million) already comes from sources other than the academy, much of it from local governments.

The institute is also increasingly involved in international projects. Apart from getting data from US satellites, the institute has made an agreement with Canada to receive data from Canada's RADARSAT satellite.

There is also close collaboration with Japan's National Space Development Agency (NASDA). The institute is receiving data from the Japanese Earth Resources Satellite-1 launched by NASDA in 1992 which has synthetic aperture radar. But because of problems with the power source of the satellite, only ten minutes of data can be received at a time.

Nevertheless, a strengthening of collaboration with Japan seems likely. A NASDA engineer is currently working at the institute and, according to Liu, NASDA is hoping to set up a receiving station for Japan's Advanced Earth Observing Satellite (ADEOS) at Urumuchi in the remote north-west of China. □

## China slowly getting wired

RESEARCHERS at the National Laboratory of Medical Molecular Biology in Beijing keep personal computers in their homes so they can use the Internet at night when the lines are less busy.

Chinese scientists are eager to use the Internet, and most scientists in leading research institutes and universities now have e-mail numbers. But the number of points of access to the Internet are still very limited, and their capacity is low.

As elsewhere in the world, it was physicists who first got China linked to computer networks of the outside world through the Institute of High Energy Physics in Beijing, which is still a key node for Internet access. But the institute's link often gets overloaded during the day, hence the need for computers at home.

The Chinese Academy of Sciences runs a local area network linking 30 of its institutes and Beijing and Qinghua universities in the Heidan district of Beijing, which is being developed as a 'science city'. The network has a fairly low-capacity international link of 64 kilobits per second through Hong Kong.

But academy scientists complain that the academy network is beset with bureaucratic problems. It is, for example, closed down at the weekends.

The ministry of post and telecommunications also runs a network with Internet

## Beijing Astronomical Observatory moves into real estate

As part of the reforms in the mid-1980s aimed at making government research organizations more responsible for running their own affairs, the Chinese Academy of Sciences offered its institutes the opportunity to buy up the land they occupy. Few took up the offer. But Beijing Astronomical Observatory grabbed the chance and has come away smiling.

"We got in just in time" says Nan Rendong of the observatory. "We bought in 1987, and the price has shot up 3 or 4 times in value". The land is an extremely valuable asset for the observatory because all research institutes and universities in China have to provide housing for their staff. Last year, the observatory built new homes for 73 families with an outlay of 4 million yuan, or about the same amount as the total operational funds it receives from the academy. The observatory receives a small amount of rent from the families and manages the real estate.

The new houses provide a strong incentive for attracting young researchers to the observatory. So keen is the observatory to get young staff, particularly from overseas, that senior staff often give up their right to larger housing space to younger people. "That is the way we

have to do it" says Nan, who gave up his option on a three-room apartment and continues to live in a small two-room ground floor apartment.

But despite the wise investment in land, the observatory is still struggling to survive. It has had to close its Tienjin observatory in

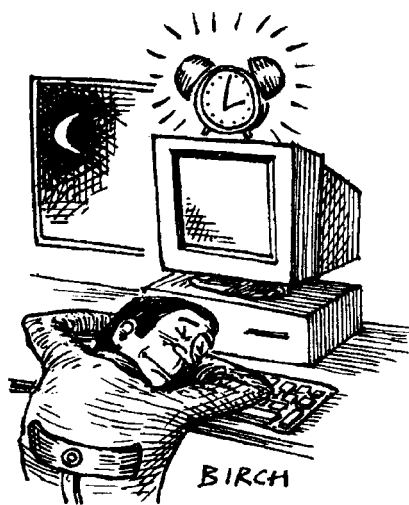


One of the 28 antennae of the metre-wave aperture synthesis radio-telescope at the Miyun Station observatory on the Miyun reservoir.

the western suburb of Tienjin and many staff have had to move to venture companies set up by the observatory. Further contraction is expected as the observatory grapples with the new market economy. □

access. But again, scientists complain that the ministry has been "very slow" to develop Internet access, although things may improve next year.

Li Maoming, deputy secretary general of the State Science and Technology Commission, says that China has an



ambitious 'Golden Bridge' project to link the country with computer networks. The State Education Commission will also soon create a national network linking 100 universities.

The network is called the Chinese Education and Research Computer Net-

work (CERNET), and Dong Weiguo, deputy director of the education commission's division of finance and planning, says it should be operational "by the end of this year".

Qinghua University will be a key node with an international link of 128 kilobits per second, again possibly through Hong Kong. There will then be a backbone linking other leading universities around the country, for example in Shanghai and Nanjing, and lower capacity branches off this.

In addition to this network, the China Medical Board in New York is funding a network linking 11 medical universities in part to help communications between researchers involved in the human genome project.

Not surprisingly, there is concern in some government circles about allowing China's citizens free access to the Internet. As things are, the general public are not allowed to buy antennas for satellite television to prevent access to foreign news programmes. And China's government-controlled newspapers now often warn of the "social dangers" of the Internet, such as the availability of pornography.

Nevertheless, it seems that government officials responsible for science policy see that access to the Internet is essential for the development of science in China, and rapid growth of networks appears imminent. □



# Universities more equal than the others

THE State Commission has picked out 36 of China's universities for special attention, giving them direct access to its own funds, but not all of those are equal in esteem. Just a dozen or so of them are by general consent outstanding, and among these, three are conspicuous: In Beijing, Peking and Tsienghua Universities, whose relationship with each other resembles that of Harvard and MIT in Cambridge, Massachusetts, or the Universities of Tokyo and Kyoto in Japan. But there is also Fudan University in Shanghai, which enjoys the benefits of distance from the centre.

What makes them special? At bottom, it is the old business of good students flocking to where there are good teachers. But these three also *look* different. Like the University of Tokyo, they are all comfortably enclosed by a wall that makes the campus seem what its name suggests, a safe haven. (Although Fudan's internal roads are crammed with bicycles, which makes them far from safe.)

Fudan has the more interesting problems. Shanghai, one senior academic says, is a unique place, with ample opportunities for development. For one thing, it is growing quickly, mostly by migration; by 2020 the population will have increased from the present 13 million (plus 3 million migrants) by an estimated 25 per cent.

Fudan has successfully made friends with the municipal government, looking to the mayor (currently a scientist, Xu Kuangzhi) not just for moral support but for funds — in return for services rendered. That side of the bargain explains what is called the Fudan Development Institute, an interdisciplinary research centre on the campus whose role is to keep an eye on the development of the Shanghai region as a whole, forecasting obvious changing patterns, but also looking out for problems and opportunities. A report last year on the development of the Yangtse delta appears to have won the government's approval.

None of that implies that Fudan is free from problems. More acutely than its rivals in Beijing, the university suffers from the gap in the age-distribution of its staff that is the enduring legacy of the Cultural Revolution. A large proportion of the teaching staff will be retiring by the end of the century.

At the same time as it will be recruiting their replacements, Fudan hopes to expand the numbers of its graduate students, now 2,500, towards the point at which graduate and undergraduate students (now 8,000) are roughly equal. That depends on an increased supply of research funds; like its rivals, Fudan believes there will be no great difficulty.

Peking University has a somewhat dif-

ferent air, but that is only what would be expected from such a stylish university. The campus has room for the occasional ornamental lake, but otherwise the business of being the best occupies the university. Nowhere, in the fight for quality, has the convention that the old know best been abandoned with such enthusiasm.

By China's standards, Peking is big, with more than 8,000 undergraduates, getting on for 4,000 graduate students and roughly 7,500 students who earn degrees by part-time study or even by correspondence courses.

But the research is what makes the university strong. Mathematics is its strong suit, supporting 150 graduate students and 15 postdoctoral students. But even pure mathematics is not all that pure in China these days; the mathematics department helps to earn its keep with the help of outside contracts.

Chemistry is the largest research school (with a faculty approaching 300 people), but that includes the national centre for the rare earths (which are less rare in China than in most places of the surface of the Earth). One of the centre's goals is to understand the small differences of the properties of the rare earth elements, another is to find novel uses for them (as in nonlinear opto-electronics).

But the hard core of the chemists' interests is understanding the mechanical properties of molecules, dynamic and otherwise. They pride themselves in publishing mostly in the *Journal of Chemical Physics* and they have all the appearances of being ready to take over that neck of woods when they are properly organized.

There are other things on the boil. For example, the National Protein Engineering Laboratory, part of the university, is busily cloning the genes expressed in the cells of the rice stalk in the hope of identifying the factors that give it strength. It has failed to persuade Japanese partners to join it in a search for ways of regenerating monocots (such as rice) from single cells, but has pushed ahead regardless, and reckons to control the world's largest company (about to be listed on China's primitive stock exchanges) for producing  $\alpha$ -interferon.

But the university's most remarkable boast is that it has earned huge sums of money by its software-hardware system for setting type in Chinese characters; the system is said to be used by the *People's Daily*, while the university is thinking of setting up a branch office in Japan. Peking will not give up these advantages lightly.

Yet Tsienghua, geographically its nearest rival, has Peking within its sights. The

rivalry would be a joke were it not serious. At least some of it goes back to when, during the Japanese war, the two universities were exiled to the southwestern city of Kunming. Then, Tsienghua was largely an engineering school, but it is now turning itself into a comprehensive university.

The university seems to have good friends. The faculty numbers 3,000, but a third of them appear to spend all their time on research. Although the humanities faculties are growing, the intellectual centre of gravity at the university remains on the engineering side.

The president of Tsienghua, X. Y. Wong, is a nuclear engineer by background, and is still proud that the university has a research reactor not far from the campus. Indeed, until a few years ago, the university was also providing domestic heating for some of its buildings by means of a helium-cooled uranium reactor operating at 200 °C. He believes there is an immediate future for nuclear district-heating schemes.

For the rest, Wong's concern is to improve the quality of the faculty. Evaluation is his watchword. He plans that laboratories and other units of the academic enterprise should be evaluated every one or two years. By that means, one infers, it should not take long to shake Peking on its pedestal. □

## Sources of error

CHINESE names have always been confusing to other people, but they have recently become more so as many Chinese have adopted Western conventions. In spoken and written Chinese, the family name comes first, followed by the given names, so that the president of the Hong Kong University of Science and Technology (see page 550) is strictly Professor WOO Chai-Wei, but his own university's literature calls him Dr Chai-Wei WOO.

In this survey, the authors have quoted names in the form in which they are given on the person's visiting card, but have then used the surname alone in subsequent references. But it is not always possible for the uninitiated to distinguish family names from given names, for which reason it is hoped that nobody referred to in the accompanying text will be offended that he or she is over-familiarly referred to.

There is a further complication in the use of the transliteration of Chinese place names. Since 1971, the capital of China has been 'Beijing' in non-Chinese languages, rather than the previous 'Peking'. But the University of Beijing now apparently prefers to be known as the University of Peking, and Qinghua University (also in Beijing) as Tsienghua University. □



## Faith in academic excellence

CHINA is inevitably a blend of the old and new, but the old may be just ten years old, when the Institute of High Energy Physics at Beijing was halfway through construction of its electron collider machine. But why does China need a particle accelerator, when economic and social problems abound?

One answer is Mr Deng Xiaoping, China's paramount leader. Deng appeared both at the ground-breaking ceremony in 1984 (when he said "I'm sure this must be right") and at the inauguration of the accelerator in 1988 (when he said that "China must have her own position in the development of the world's science and technology").

The laboratory has more than that to say. Although the energy of the colliding electron-streams is only 2 GeV, or rather less than the energy of electrons in modern synchrotrons built as light-sources, it has by all accounts proved to be a well-designed machine and also a prolific source of B-mesons.

That has set the laboratory brooding about the construction of a 'B-meson factory', or a collider optimized to produce B-mesons. The laboratory already has the funds for a feasibility study, and hopes for a decision of the outcome in about eighteen months. Deng's words will be much quoted in the application.

But two weeks in China confirms that this is not where the action is. The young bright sparks, especially those returned from spells abroad, are working at the intersection between physical science, molecular biology and biotechnology or are seeing what can be done with the surface of silicon, trying to build nanostructures or otherwise to manipulate them.

One such is Professor Chen Zhu, who



The particle accelerator — Deng Xiaoping's support may build the next.

works under the label of haematologist at Shanghai's Medical University No. 2. Chen and his wife, Dr Sai-Jua Chen, send their 11-year-old son to a boarding school in Shanghai so that they can have more time in the laboratory. "Perhaps we're not very good parents", Chen says. On a Saturday in October, the son returned from school for the weekend and was doing his homework at the bench while his parents carried on with their work.

Chen, who was elected to the Chinese Academy at the end of October at the tender age of 42, has one important feather in his cap — a study of acute T-cell leukaemia (of which there are more than 50,000 cases a year in China) showing that the proximate cause of the disease is an 11:17 translocation in the genome of patients, which in turn disrupts the gene for a known retinoic acid receptor.

The two Chens have gone on from that to argue that arsenic (which induces apoptosis) may be a cure for the leukaemia; the first successful treatment in Shanghai was in 1985, but the remedy is now being tried out systematically at Harbin Medical University in northern China.

The same observations have led to the identification of three novel genes believed to be effective on differentiation.

Chen now plans to do the job more systematically, for which reason he has acquired a set of Daniel Cohen's YAC clones from Paris, putting them at the disposal of China's strapping Human Genome Project, supported so far with a grant of only 3 million yuan from the National Natural Sciences Foundation.

So far, 15 laboratories are participating in the project, in which the Academy of Medical Sciences will play a central part. The project is anxious to make a start on the genetic analysis of human diversity, on the grounds that China is



The Rockefeller Foundation's Union Medical College of the 1920s, now the Academy of Medical Sciences.

changing so quickly that present diversity born of genetic isolation may soon be lost.

Chen himself has meanwhile signed a contract with SmithKline Beecham to make a study of the 'expressed sequence tags' (ESTs) to be found in the differentiating cells of the haematopoietic system; he has until the end of 1996 to complete the job.

The heartening circumstance, for China, is that Chen is by no means alone among younger scientists. Like him, they are both ingenious and hardworking. They take it as a given that they will publish in international journals, although *Nature* is surprised how little demand for space from China there has been. □



The newly opened botanical garden of the Institute of Botany, Guangzhou.



## Academics plump for business

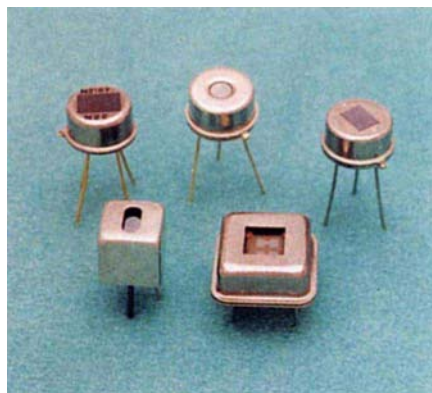
ALL research institutes are angling after joint ventures with external industrial enterprises, but the Shanghai Institute of Technical Physics has been luckier than most. For just over a decade, the institute, formally a dependant of the Chinese Academy of Sciences, has been in partnership with the Nippon Ceramic Company Ltd. The tail is not yet wagging the dog, but it has become a big business.

The institute's main expertise is in infrared physics, to which it graduated in the mid-1960s, no doubt in step with growing military interests in this field. It has become skilled in making infrared detectors and filters from synthetic materials such as indium antimonide (InSb) and the like.

The Japanese collaboration, called Shanghai Nicera Sensorco, is a joint venture of the kind allowed between Chinese and external companies: the equity is shared equally between the partners, profits qualify for a reduction on the normal corporation tax but the agreement has a limited duration (in this case 20 years), which may be thought China's way of keeping its options open.

The company began by making infrared filters, which the institute proper could make only in small numbers. The initial investment (of 3.2 million yuan) was provided equally by the partners, but in 1988 the Japanese partner increased its investment by buying land and putting a building on it; an extra storey was added before construction was complete.

The joint venture quickly added infrared sensors (for which there are endless markets in such things as automatic



Not transistors but infrared detectors.

doors) to its repertoire, and now produces them at a rate of several millions every year. The total volume of sales from the venture (handled in parallel by the Chinese and Japanese partners) was 130 million yuan in 1994, of which the profit amounted to 42 million yuan.

But China gets more from the venture than the share of the profit kept by the Institute of Technical Physics. From the outset, it was agreed that the general manager of the enterprise would be Japanese (with a Chinese physicist serving as chairman of the board), and that production would be organized on familiar Japanese lines. The deputy general manager says that he and the workforce of 200 have had to adapt to Japanese ways of punctuality and of keeping clean-rooms clean, but that everybody is now reconciled to the regime.

The export business seems to be substantial. The company reckons to have 60 per cent of the world market in infrared detectors. With 85 per cent of all its products manufactured in China, the economic benefits cannot be negligible.

Yet relations with the parent institute remain friendly. The company's first Chinese staff were recruited from there (and retained the right to return if something went wrong). Since then, the company has stepped up its research and development effort in order to win market share in parts of the world where its market share is not strong, as well as to develop new products to keep ahead of its rivals.

The result has been the further recruitment of institute staff (this time without the guarantee of returning) and even the hiring of laboratory space in which to house the research and development department. With business growing by 30 per cent a year, they may need a new building soon.

Meanwhile, the institute's

entrepreneurial instincts appear only to have been stimulated by this success. It is the sole owner of another company in the business of contracting for major infrared projects of various kinds, and seems on the lookout for other opportunities. One of the benefits of this way of looking at things is that the regular staff of the institute have become more mobile than they used to be. □

## A living legend

CHINA's best known computer company was called Legend even before it became one. Just over ten years old, it boasts that its average annual rate of growth over the past decade is 122 per cent. Total sales amounted to 4 billion yuan. In the best traditions of computer companies, it began life in two small rooms (owned, in this case, by the Chinese Academy of Sciences).

Legend's first claim on market attention was its software for displaying and then printing Chinese characters, but since 1984 it has won a reputation for designing the electronic environment of microchips so as to squeeze the most out of their performance. The result is that its staple products are computer motherboards designed and manufactured to meet other computer-makers' needs. Close on their heels is a range of add-on or plug-in peripherals.

But Legend also puts together personal computers, buying Pentium chips from Intel when it needs them. Last year, Legend says, it sold 45,000 of its computers at prices lower than its competitors in the West could make equivalent products. The numbers would not worry the IBMs or Compaqs of this world, but Legend's global ambitions just might.

Unlike most Chinese enterprises, Legend has set about its expansion overseas by the methods followed by multinational companies elsewhere, which means setting up companies in different places and controlling them from Beijing. There has been a Legend Company in Hong Kong since 1988, but the company now has sales outlets (often just backstreet shops) in places as different as Los Angeles and London.

But the company has also retained a substantial systems analysis function of the kind with which it paid its way in the world when it was much smaller. One of its odder current software contracts is a scheme for integrating the people-databases obtained at local police stations when local residents register their names on moving residence (as required by law) into a hierarchical system that will be effective at the provincial and even national level. One can see the attractions easily; Chinese seem insensitive to the drawbacks. □



More than a million devices a year.



# When is prenatal diagnosis 'eugenics'?

'EUGENICS' seems to have become a forbidden word, or at least one that causes great embarrassment. A survey of China in 1995 would manifestly be incomplete without some explanation of the great fuss earlier this year when the official Chinese news agency reported the adoption of legislation aimed at the improvement of China's genetic stock.

Most people in China are either genuinely confused about what has happened, or are anxious to spare the feelings of innocent visitors from elsewhere. But curiously, when a group of senior academics at Fudan University in Shanghai said that the one "who would know about that" was the demographer in charge of the university's population unit, the man himself said, "Oh no! There is no law, but only a regulation applying only to a single province in the northwest."

The Academy of Medical Sciences at Beijing seemed a more promising place

at which to enquire, but there the answer was that the Institute of Genetics would have more to say on the subject. Fortunately that was the case. Dr Souyi Chen, the newly appointed director of the institute, produced not merely an English translation of the law, but copies of letters she had written to colleagues elsewhere in stout defence of the location of the next (18th) International Genetics Congress at Beijing in 1998.

The law is described as one affecting "Maternal and Infant Health Care" and was adopted by the Eighth People's Congress on 27 October 1994, more than a year ago. What seems to have happened this summer is that the law came into force on 1 June 1995, and that the news agency took that as an opportunity to publicize the fact.

As early as March this year, doubt was cast on the future of the 18th Genetics Congress by a fax from Professor A. J. F. Griffiths, secretary-general of the International Genetics Federation (IGF),

polling members of the executive committee to choose between shifting the venue (Australia had been mentioned separately) and insisting that there should be a full discussion of the new genetics law at Beijing.

In a hot-tempered reply, Professor Li Zhensheng, the president of the Genetic Society of China, says that the proposed poll was unconstitutional and "therefore groundless", that it offended against the statutes of the IGF, which specify that it shall not "attempt to influence legislation", and that the legislation did not in any case entail compulsion.

On the last point, Li quotes article 10 of the law (see box) and says that "it is clearly stated that measures can be taken only if both the male and the female agree, and no punishment can be enforced if they insist on having babies". But this is the article that makes sterilization the condition of a certificate without which marriage is forbidden.

In her letter to Griffiths, Shouyi Chen, writing as secretary-general of the 18th Congress, accuses him and Professor D. A. Smith (from the University of Birmingham, UK) of using surreptitious means of seeking to move the congress to Australia, says that the new law in China only facilitates practices common "for decades... in western countries" and says that there is no similarity between what is proposed in China and "Hitler's concept of eugenics".

Chen is not nearly as fierce in the flesh as she is in prose. In the event, the venue of the 18th Genetics Congress will remain in Beijing. Chen's view of the law is that it is simply a step towards the provision of the antenatal diagnosis already readily available in the West.

Chen is a plant geneticist who is hoping that genome sequencing will eventually bring to China an understanding of the rice genome. Her group is already working on the identification of important genes in rice, having identified a fungal resistance gene, but she hopes there will be a more concerted effort in the field, in which her institute is planning to take part in a substantial project. To be fair, the law is also replete with a specification of the conditions that must be satisfied by those administering the maternal and infant health-care programme. Institutions, for example, are enjoined to raise the "medical technical level" and to "take all measures for the convenience of the people". Those engaged in genetic diagnosis have first to pass an examination, as must home midwives, and then be certified to carry out the work. The law also says that all personnel "shall strictly abide by professional ethics and keep the secrets of the individuals concerned".

## The Eugenic Law

THE provisions of the law on maternal and infant health-care that may elsewhere be regarded as draconian are as follows (where the bold numerals refer to the article of the text):

The purpose of the legislation is to ensure the health of mothers and infants and to improve "the quality of the newborn population" (1), to which end "people's governments at all levels" shall take the lead (3), subject to guidelines from the national Health Department (4) specifying what principles should be followed in different areas of China. The state undertakes to support research and education in the field (5).

In the chapter dealing with premarital health-care, the law specifies that the services to be provided include "education in sex health, reproduction and genetic diseases" and a premarital "physical check-up" for both partners "to see whether they suffer from any disease that may have an adverse effect on marriage and child-bearing" (7).

Physicians are to look for "(i) genetic disease of a serious nature ... [and] (iii) relevant mental diseases, issuing the couple with a certificate afterwards (8). Physicians have a responsibility to "explain and give medical advice" about any adverse medical findings; the law says that the "two may be married only if both sides

agree to take long-term contraceptive measures or to take the ligation operation for sterility" (10).

There is a right to a second opinion (11), but otherwise a couple must produce their check-up report "when applying for marriage registration" (12). Local and provincial governments will determine what charge should be paid (13).

Elsewhere, the law says that "if a physician detects or suspects" that a married couple "suffer from genetic disease of a serious nature", he or she shall give medical advice to the couple", who are then required "to take measures in accordance with the physician's advice" (16). Physicians are also required to advise a termination of pregnancy if, at a pre-natal diagnosis, they find that the fetus has a "genetic disease of a serious nature" or a "defect of a serious nature" (18).

Interestingly, the law also says that "sex identification of a fetus by technical means shall be strictly forbidden" (32). "Genetic diseases of a serious nature" are defined as diseases caused by genetic factors that may totally or partially deprive the victim of the ability to live independently, that are highly likely to recur in generations to come and that are "medically considered inappropriate for reproduction". "Relevant mental diseases" refer to "schizophrenia, manic-depressive psychosis and other mental diseases of a serious nature" (38). □



## Hong Kong's boom in education

**Hong Kong** likes to think of itself as the jewel in mainland China's crown, and it looks increasingly like a jewel. Everywhere there is splendid architecture. Last decade's most admired building by Richard Rodgers, the headquarters of the Hong Kong and Shanghai Bank, has been almost literally put in the shade by Pei's structure of blue reflecting glass designed to give the illusion that, from all directions, it is simply a planar sheet reaching for the sky.

Why all this investment in real estate almost on the eve of the lapsing of the British lease on Hong Kong on 1 July 1997? The simple answer is that Hong Kong believes the Basic Law (the *anschluss* treaty between Britain and China) will be faithfully honoured. That means that there will be a moratorium on changes in the way Hong Kong is run for at least the next 50 years. "One country but two systems" is mainland China's slogan for the future.

The more complicated answer is that the continuing activity is a way of whistling to keep the spirits up. Already, China has announced that it will change the way in which Hong Kong's Legislative Council is elected. In the past few weeks, there has been trouble over the contents of a Bill of Rights adopted by the Hong Kong government, too late in

the day in China's view. The opportunities for influencing events informally will be greater after 1997.

The truth is that there is an undertow of anxiety, especially among academics. Some plan to leave at the end of this academic year. Others note wryly that they still have their 'passport', by which they mean the right to live and work elsewhere.

But there are many more who have recently migrated from the United States or Taiwan, saying in effect that they believe Hong Kong to be the ideal vantage point from which to witness not just the greatest firework display ever (planned for the fateful 1 July) but a unique social and political event. The occasion will not mark the winning of independence from colonial tutelage (Britain's) but the transfer of one small country from a relatively liberal regime to one that is less so.

Nevertheless, the academic sector in Hong Kong seems to participate fully in the boom that dominates the rest of the city. The government is proud of having increased the participation rate in higher education to about 18 per cent (to which, it says, must be added the 5 per cent or so of young people who are educated elsewhere). That trend will continue at least for another 18 months. □

## Building an MIT in just seven years

HONG Kong's newest university, the Hong Kong University of Science and Technology (HKUST), owes its existence to what is called the Jockey Club, constituted as part-charity, part racetrack operator, which in turn owes its financial muscle to the legendary willingness of Hong Kong's 3.5 million permanent residents to bet on horses (and most other things that move or can be rolled on green baize surfaces).

The university is out of sight of Hong Kong's several downtown areas, but instead on Clearwater Bay, a delectable stretch of ocean opposite an island a few kilometres away. There are a few ships at

anchor on a jetty there; the president of the university, Professor Chai-Wei Woo, apologetically promises that a scar on the centre of the distant island will be healed when it becomes a golf course.

Woo (like most of the colleagues he has recruited) is a Chinese-American. A physicist, he was president of the San Francisco State University at his recruitment at the end of 1987. He has described how he and his wife set about creating a faculty for HKUST with the help of a fax machine in their kitchen. The result is a star-studded group of teachers, one of whom says: "Now that we're all here, we will have to show that we can work."

So far, it has all been somewhat breathless. Construction began late in 1989 and doors opened to the first students (roughly 600 of them) two years later, in 1991. But the university plans to have 7,000 students in October next year, nine years after it was conceived. In his address to the convocation (degree-giving ceremony) on 3 December this year, Woo was lyrical about the speed with which the development had been



Gateway to technology.

carried through: going "from ground-zero to full-size in five years flat".

The Jockey Club seems to have paid for the speed. Its original offer of HK\$1.6 billion, matching a similar offer from the government, had to be increased to HK\$1.9 billion to accommodate the "fast-track" construction system Woo is proud of; that is a euphemism for ensuring that design is always ahead and not behind construction. "Everything had to be done at the same time", he says.

In his address, Woo also promised that the faculty would soon be able to relax. For the time being, next year's student target will be a kind of plateau. Teachers will be able to cut back on administration, and instead volunteer for what might be called pastoral duties, helping students to broaden their outlook on life in general.

He is proud of running a tight ship. Under a scheme introduced in 1993, academic performance is reviewed at regular intervals and, sometimes, the result is the "withholding of salary increases and the termination of contracts". There is a plan to operate a similar scheme for administrative staff, but with the difference that clients will be members of the review panels. For students, there is continuous assessment.

The managerial ethos also entails operating within budget. Woo says that the university has returned money to the Hong Kong government in each year of its operation. Part of the explanation is that HKUST has successfully dipped its feet in business.

It reckons to be the most successful Internet provider in Hong Kong under the trade name of Supernet, and formed a joint venture in October with one company in Singapore and two in Japan to form an integrated Asian venture. The university also boasts of a share in a HK\$120 million contract to provide local weather warning systems for Kong Kong's new airport. And these are only early days.

When the student target of 7,000 is reached next year, growth of the university will depend in part on HKUST's capacity to capture able graduate students; a final target of about 10,000 students has been agreed. Woo is looking to the north, and to Shanghai in partic-



Hong Kong University campus. The view from Clearwater Bay.

ular, for them. (He will not be surprised that his hospitality is taken as a sign of weakness; northern colleagues say that HKUST must have a deficiency of good graduate students.)

Meanwhile, the university seems bent on following the practice of universities in mainland China of augmenting its income by playing the market; the habit of the north is infectious. Its technical interests are concentrated in computer software and hardware, materials science (including optoelectronic devices) and telecommunications. There are nearly a dozen research institutes, including one in telecommunications and another in biotechnology. But the theory of business also looms large; there are MBA courses and departments of marketing.

HKUST takes a business-like approach to patenting innovations in its research departments, trading off the high cost of securing protection against a hard-headed estimate of the likely income from its possession of a patent. Its portfolio may one day include patents for a neat way of making diamond films by the ablation of polymers with laser radiation in the presence of particular catalysts and a solid-state magnetometer of high sensitivity that promises to be a novel means of reading the information on computer hard-disks.

How does this technological university, which openly compares itself with Massachusetts Institute of Technology, see its place in China after 1997? Woo has no doubts. He wants the university to be "a catalyst for change not only in China but in the Western Pacific rim as a whole". Already there is an association of Pacific technological universities (including places such as South Korea). But more than any of Hong Kong's other universities, HKUST is likely to be the most immediately engaged with China.

How does the faculty regard the prospect? One senior academic says that he was drawn to HKUST eighteen months ago by the "attractions of a blank slate", another that "this is the first time that I have made a decision [about the direction of my career]". But yet another, after a brief spell in Hong Kong, plans to go back to the United States at the end of this academic year. "At this stage in my life, why should I put up with the hassle?", he asks. □



**Chia-Wei Woo, academic entrepreneur.**

## A British academic outpost in change

EVEN the oldest university in Hong Kong is bustling with construction cranes, cramming extra living accommodation onto the crowded hillside where it has been since its foundation in 1911. Like many others, Hong Kong University (HKU) grew outwards from a medical school.

For much of its history, the university has been a British Commonwealth university in the traditional mould, with the core of the academic body trained and recruited in Britain. But that has changed dramatically in 20 years. HKU, with 11,000 students one of Hong Kong big three, is as eager to see what the new world has to offer as the more overtly Chinese universities.

The student body is more cosmopolitan than at the other Hong Kong universities, with students from 46 different countries (including Belgium and Hungary). More than 2,000 are from mainland China.

The Commonwealth tradition lives on most vividly in the requirements for entry to the university (as to the others in Hong Kong). Entry is determined by a student's grades in school-leaving examinations called 'A levels', patterned on a British series of examinations. As in Britain, so in Hong Kong, academics are beginning to wonder whether the system may not be too narrow a basis for higher education. There are also problems of literacy; Chinese-speaking students (who speak Cantonese and not Mandarin, as in the north) often require remedial classes before they can write their language confidently.

Although the university is the most comprehensive of Hong Kong's seven, teaching a catholic range of courses, it has offered research degrees for only the past three decades.

The immediate goal is to strengthen research across the board, to which end the university is pleased that it now has \$HK10 million a year to employ up to 25 post-doctoral fellows (previously non-existent) on its own initiative. (Others come from overseas, bringing their own funds with them, or are supported by research grants.)

As in China proper, graduate students are paid a stipend (of about \$HK140,000 a year) by the university out of its recurrent funds, while the ubiquitous Jockey Club has provided the grant with which to build a new building (now under construction) in which some of them will live.

The university's view of its future within China is variegated. The general opinion is that it will be China's academic window on the outside world, and on Europe in particular. But others point to Hong Kong's rich endowment in research equipment and its sheer richness: China, people say, cannot afford research projects costing \$HK1 million, but Hong Kong can.

Already there are large numbers of collaborations with academics in mainland

China, according to Professor Brian Weatherall, director of the School of Research Studies, but these are mostly undocumented. By 1997, the chances are that collaborations with people to the north will be the rule rather than the exception.

There are also particular ways in which the Hong Kong academics reckon that they can help stiffen China's academic work. One of mainland China's weaknesses is social science, according to Professor John Bauer-Shore, the statistician who is head of the department of social sciences at HKU. He says that the "Chinese do not trust even their own data".

In fields where conceptual matters are more important, he says, the caste of mind of Chinese academics has been cramped by decades of ideological conformity. Evidently the years ahead offer great opportunities for the proselytizers of social science.

Academic opinions about 1997 are as elsewhere. Some are plainly in Hong Kong for the interest of what lies ahead, but others are keeping an open mind. □

### Plagiarism dispute ends

THE University of Hong Kong is relieved that a long-running legal dispute between two of its academics appears finally to have ended. The dispute has centred on allegations by Linda Chih Ling Koo, a lecturer at the university, that Professor Tai-hing Lam from the medical school had plagiarized a questionnaire that she had prepared as part of an investigation of the reasons why Chinese women appear to be particularly susceptible to lung cancer caused by smoking cigarettes.

The allegations were accepted as true by the lower courts in 1992, which judgement was endorsed by the Appeal Court in Hong Kong. The university then set up a review to decide Professor Lam's future in the university; the report of that inquiry, in July this year, decided that the circumstances of the case were not such as to require his removal from the university staff, whereupon Dr Koo applied to the Supreme Court of Hong Kong for judicial review of the university's decision.

The court has now (on 24 October this year) refused that application on the grounds that the judge (Mr Justice Findlay) had not been persuaded that the conduct of the university's review of the case "did not breach the duty of fairness to the applicants". Costs were awarded against Dr Koo.

And the university's response? Professor Brian Weatherhead, the vice-president for research, says he is writing a document to guide academics at the university on the importance of plagiarism as a form of academic misconduct. He does not say when it will be published. □



## Prosperity and the Basic Law

HONG Kong's expansion of higher education since 1988 has been deliberate. The objectives are also clear: to increase the participation of Hong Kong's young people in higher education so as to increase the territory's supply of modern skill. But why bother, with 1997 just 18 months away? Much the same might be asked of the huge investment planned in Hong Kong's new international airport, which is unlikely to open for business until the end of the century.

The explanation is simple. The Basic Law defining what will happen after 1997 is explicit on the arrangements that will apply when Hong Kong becomes a Special Administrative Region of China. The territory will be then what it is now, financially autonomous, and will remain so for at least 50 years. The government's finances will remain its own, to be used "exclusively for its own purposes". The Law says that they shall "not be handed over" to the government of the People's Republic of China.

If that is indeed what happens, Hong Kong has every interest in enlarging the numbers of young people qualified at modern jobs. That, no doubt, is part of the reason why Hong Kong's academics have also been dealt with generously over the years.

Salaries appear ample. The pay of a lecturer ranges upwards from HK\$25,000 a month to HK\$57,000 a month, but the average salary of a professor is HK\$94,000 a month and may be much greater if he or she shoulders extra responsibility. These are salaries on an internationally respectable scale, abated though they may be by the high cost of living. They provide the universities with both a motive and an excuse for strict internal evaluation procedures with the objective of making sure that dead wood does not accumulate.

But there is general agreement that Hong Kong differs from the other 'tigers' of East Asia (Japan, South Korea and so on) in spending less generously on research and development. During the past three decades of growth, Hong Kong has moved away from manufacturing clothing and electronic consumer goods cheaply towards telecommunications on the one hand and the provision of financial services on the other, neither of which is dependent on indigenous technology.

It seems now to be recognized that the balance will have to be redressed; a speech to this effect by Dr Jeremy Bray, a member of the British Parliament, in late October seems to have left a deep impression.

That explains the Hong Kong government's repeated attempts at new grants schemes for the support of research. Last year, for example, the Department of Industry set up a scheme for distributing

HK\$300 million a year for research and development projects likely to strengthen Hong Kong's manufacturing industry. The rules say that applicants will usually be research institutions and universities, but that individual companies may occasionally apply.

For basic research within the universities, there has been since 1991 the beginnings of a competitive research grants system, still administratively embedded within the University and Polytechnic Grants Committee (which provides public funds for all the university institutions and which traditionally has included an allowance for research). By Hong Kong standards, spending is still modest (HK\$150 million last academic year), but there are signs that this will improve (per-

haps at the expense of the universities' recurrent budgets).

Whether Hong Kong will succeed by these means in reconvert itself into a manufacturing centre is another matter. Like Singapore, its scope is limited by the number of its permanent residents (3.5 million). The most likely course of events is that there will be a string of agreements with companies based in the free economic zone of Shenzhen, just north of the territory (and in which Hong Kong businessmen already have important financial interests).

Plainly, everything will depend on how well the Basic Law protects Hong Kong from the inevitable envy of China proper. The strongest case for hoping that 1997 comes quickly is that only then will it be possible to tell for sure. □

## Hong Kong's one-topic science park

NOT everything the Royal Hong Kong Jockey Club touches turns to gold. So much is clear at the Institute of Biotechnology (HKIB), on the campus of the Chinese University of Hong Kong, where the handsome building built in 1989 with the help of \$HK170 million of the charity's funds is still only partly occupied.

Dr Albert Y. Chang, the director, is not too downcast. And there is no question that his venture is an unusual experiment. The institute is founded as a company in its own right, but with a mission to stimulate the development of biotechnology in Hong Kong. It is constituted not as if it were a laboratory but as a one-building single-issue science park; that it is still looking for a full complement of tenants is in no sense its own fault.

How will it make a living for itself? One possibility is a scheme to become the home of small companies making their way in the new world of biotechnology. In the Chinese style, there is a name for these activities: the 'Incubator Programme'. Floor space for offices and laboratories is cheap as Hong Kong prices go (\$HK35-\$HK40 per square foot per month), in return for which occupants also have access to common facilities (such as dark rooms, laboratory equipment and even parking space). Tenants can replace up to half their monthly fees by equity in themselves.

But the laboratory (like other Hong Kong companies) can also apply for research and development funds from the Hong Kong government's Applied Research Fund as well as other private sources. This is the spirit in which, last year, the institute embarked on a scheme for the mass culture of cells from the ginseng plant (East Asia's equivalent of cola drinks). The hope is that the present venture with a 20-litre bioreactor will have led to the development of a full-scale industrial process by late in 1997, a little after the fireworks display on 1 July.

There is also a substantial hope that the institute will be selected by the World

Health Organization as the contractor to manufacture for impending clinical trials in China quantities of the recombinant malaria vaccine developed at the US National Institutes of Health by Dr David Kaslow.

Meanwhile, the institute has established (in 1990) a joint venture with Syntex



HKIB as an incubator facility.

Pharmaceuticals International to screen natural materials indigenous to the region for potential drugs. The idea is that the company and the institute will jointly fund the development of drugs emerging from the programme, and similarly share the profits.

The difficulty with these arrangements is that the institute itself has only a minimal core of expertise in basic research. The obvious analogy is with the Institute of Cell and Molecular Biology at Singapore where the strategy is to use a very large group of basic researchers as the potential source of spin-off companies (but where there is also a partnership with a multinational drug company, Glaxo Wellcome, to search for new drugs in indigenous plants and other species).

Unfortunately, it is not self-evident where the people or the funds for such a conversion would come from. (The Jockey Club is happier building buildings that paying recurrent costs.) And there is competition even in Hong Kong itself; HKUST, for example, has a biotechnology institute of its own. □

# Contention over growth promoters

Unwelcome technical findings can cause ructions when they are made public, for which reason the technical community has a responsibility to its findings to argue its case in advance with their opponents.

**Brussels.** Suppose you hold a conference on a contentious subject and seek to persuade the world that whatever conclusion you reach, it should be taken seriously. There are several stratagems that can be followed. One is to alert the press and television people to what is afoot, in the hope that they will amplify the message. Another is to print the message on glossy paper and distribute that to the people called opinion-formers. The trouble is that neither method can be guaranteed success. Journalists may choose to look the other way, or may be forced to do so by unexpected events, while the great and the good may have other things on their minds.

That may be part of the reason why Dr Franz Fischler, agriculture commissioner at the European Commission, asked that last week's conference on growth promoters in meat production should be attended not only by 85 or so scientists in the field but also by nearly twice as many people with a secular interest in the subject — members of the European Parliament, of the European Council's Economic and Social Committee and representatives of pressure groups that legitimately attempt to lobby the European Commission on consumer issues and related matters.

The outcome would have been amusing had it not also been depressing. Is it no longer possible for there to be a public statement of an inference based on an examination of existing data without exciting the complaint that the scientific profession has risen above itself, seeking to force ill-formed opinions down other people's throats?

On this occasion, the inferences were not all that remarkable, anyway. The underlying question is whether farmers should be allowed to use substances that promote the growth of livestock, and if so, under what conditions (see *Nature* 377, 381; 1995). The issue has been alive since 1988, when the European Community (as it then was) restricted the use of anabolic agents in animal husbandry. Many farmers grumbled and some took to the illicit use of, for example, anabolic steroids such as testosterone.

The big change since 1988 is that there is now a whole battery of growth-promoting agents on, or just over, the horizon. That is not surprising. Apart from antibiotics (which reduce the metabolic load of microflora in an animal's gut) and preparations of natural reproductive hormones and their relatives (sometimes synthetic) of the kind illicitly used by athletes to enhance

muscle-power, there is a family of synthetic chemicals called  $\beta_2$ -agonists, essentially analogues of adrenaline. They appear to have been first used in the treatment of horses suffering from chronic pneumonia, as adrenaline itself is sometimes used in people; increased growth appears to have been observed adventitiously.

But that is only the beginning of a lengthening list. Why not use growth hormones proper, perhaps bioengineered, to make animals grow more quickly, or materials such as somatotrophin, which may or may not (according to circumstances) have similar effects? There is even talk of suppressing the effects of metabolic enzymes by means of antibodies, or of adjusting the periodic exposure of animals to light, which may (but may not) affect growth through the intermediary of melatonin. And there is always, of course, gene transplantation.

Which, if any, of these routes can be safely used to promote the growth of farm animals? That cannot be a simple question, but entails an exercise in risk analysis. First, the hazards (in the sense of potential damage to the health of the animals themselves and of people who eat the meat) must be identified by the standard process of imaginative scepticism. Are children, or pregnant women, especially prone to damage? Can oestrogens added to the environment cause long-term reproductive damage? And then it is necessary to search the literature for evidence for some numerical link between the doses of these materials likely to be administered to animals or unwittingly ingested by people and the risk that individuals will be damaged. That textbook statement is impeccable; the difficulties arise in dealing with particular cases.

What last week's conference decided is that there is now no reason to suppose that the use of the reproductive steroid hormones, sanctioned for more than 30 years in the United States, Canada, Australia and New Zealand, is damaging to meat-eating consumers, although there may still be room for doubt about the effects of these materials on the animals to which they are administered, notably in the possibility that they may affect behaviour.

But the  $\beta_2$ -agonists are in a different case. There has been a handful of incidents in the past few years in which numbers of people have been acutely injured by eating liver from animals treated with them. And there is too little information about the functioning of apparently simple hormones such as growth hormone for a judgement to

be made (although bovine growth hormone is sanctioned for use in Australia).

Thus stated, these findings seem innocuous enough, but the first, as everybody was well aware, fails to justify the restrictions placed on the use of steroid hormones by the European Union (EU) in the late 1980s. But the participants were also aware (they were reminded of that often enough) that decisions about the regulation of substances in the food chain may properly involve political, economic and even idiosyncratic considerations — people may prefer not to eat residues of chemicals not naturally present and, if they are in a majority, may carry the day. So the two findings were presented as strictly technical findings, not in themselves a sufficient basis for changing EU policy on steroid hormones as growth promoters.

And then the skies fell in. One speaker complained that the pretence that it possible to consider technical issues in isolation from broader issues is an abnegation of social responsibility and was the reason why Frankenstein (but he may have meant Hitler) was able to flourish. Another held that a bunch of scientists had no business telling the larger community how to conduct its affairs. The question of "Why are we talking about this issue when Europe has too much meat?" arose in several forms, and apparently in oblivion of threats by the United States that it will use international trading mechanisms (now the World Trade Organization) to win compensation for its beef-producers on account of the EU policy. Another refrain was that opinion polls have shown that EU meat-eaters prefer their beef without added growth promoters.

The depressing feature of the proceedings is the huge gulf that appeared between the technical people and the politicians; the two groups seemed to be talking about different topics. So was it a mistake to mix two such disparate groups together? The truth is the opposite; if the gulf is so wide, surely it is in the public interest that there should be more talk, not less. But is not that a waste of time? It is only necessary to imagine that the subject had been something really serious, like the use of nuclear weapons, to know that the answer must be "No!". But steroids in meat is also an important issue. The man who mentioned Frankenstein was wrong in what he said, but he might not have spoken as he did if technical people had talked to him more often. That is the lesson. **John Maddox**



# Power laws and universality

H. Eugene Stanley

SIMPLE interactions sometimes lead to complex results. For example, dynamical systems with as few as three independent variables show chaotic behaviour that is, in any practical sense, unpredictable (at least for long times). Conversely, complex interactions can lead to simple results. This is particularly true for systems formed of many interacting subunits<sup>1</sup>, such as liquids, solids or gases, which are poised at a 'critical point' where two or more macroscopic phases become indistinguishable. At a critical point, many of the precise details of the interactions between constituent subunits play virtually no role whatsoever in determining the bulk properties of the system.

To understand these surprising phenomena (known as 'critical' phenomena), 25 years ago scientists developed the twin concepts of 'scaling' and 'universality'. Systems near critical points exhibit self-similar properties; thus, to some degree, they are invariant under transformations of scale. This is the property of scaling. The word universality connotes the fact that quite disparate systems behave in a remarkably similar fashion near their respective critical points — because what matters most is not the details of the microscopic interactions but rather the nature of the 'paths along which order is propagated' from one subunit to another distant subunit.

The development of these ideas was made possible by a remarkable combination of experiment and phenomenological theory, combined with detailed study of a simple mathematical model — the Ising model. Most of the principal ideas that emerged were tested on this model system. Despite the tremendous wealth of experimental work on systems near their critical points, however, the equation of state predicted by the theory had never been tested with sufficient care. Such a test is the subject of the paper by C. H. Back and collaborators appearing on page 557 of this issue<sup>2</sup>.

In the Ising model, classical spins localized on the sites of a lattice can point either 'up' or 'down'. The ferromagnetic Ising interaction is particularly simple: if two neighbouring spins are parallel (both up or both down), then there is a negative contribution to the energy. Hence the system has minimum energy when all the spins in the system are parallel. When the lattice is in equilibrium at a temperature  $T$ , it does not reside in this lowest energy state, but explores higher energy states. The Ising model can be regarded as a crude model of a ferromagnet if we think of the classical spins as representing the constituent microscopic moments that make up the ferromagnet.

Studies of the Ising model reveal a remarkable feature. If one tunes a 'control parameter' — the temperature  $T$  — then one finds that at a certain critical value  $T_c$  spins remarkably far apart have orientations that are strongly correlated. Such correlations do not lend themselves to a ready explanation. Normally in physics modelling you get out what you put in, yet in the case of the Ising model, we 'put in' an interaction that extends a finite distance, and magically, we 'get out' a correlation that spreads an infinite distance. How does this happen?

Our intuition tells us that the correlation  $C(r)$  between subunits separated by a distance  $r$  should decay exponentially with  $r$ , for the same reason the value of money stored in one's mattress decays exponentially with time (each year it loses a constant fraction of its worth). Thus  $C(r) = \exp\{-r/\xi\}$ , where  $\xi$  is termed the correlation length — the characteristic length scale above which the correlation function is negligibly small.

Experiments and also calculations on mathematical models confirm that correlations do indeed decay exponentially so long as the system is not exactly at its critical point. At  $T_c$ , the rapid exponential decay turns into a long-range power-law decay of the form  $C(r) = r^{-\eta}$  where  $\eta$  is called a critical exponent<sup>3</sup>. If correlations decay with a power-law form, we say the system is 'scale free', because there is no characteristic scale associated with a simple power law.

Critical exponents, such as  $\eta$ , are found empirically to depend most strongly upon the system dimension and on the general symmetry properties of the constituent subunits, and not on other details of the system under investigation. We understand power-law decay as arising primarily from the multiplicity of interaction paths that connect two spins in dimensions larger than one<sup>3</sup>. Exact enumeration methods take into account exactly the contributions of such paths, up to a maximum length that depends on the strength of the computer used and the patience of the investigator. To obtain quantitative results, the hierarchy of exact results is extrapolated to infinite order. Roughly stated, although the correlation along each path decreases exponentially with the length of the path, the number of such paths increases exponentially. The 'gently decaying' power-law correlation emerges as the victor in this competition between the two warring exponential effects.

Cyril Domb, Michael Fisher and their colleagues pioneered such exact enumeration approaches in the 1970s. Among the results emerging from their efforts was a

complete calculation of the scaling equation of state of the Ising model in a magnetic field<sup>4</sup>. It is this equation of state that is tested by the experiments of Back and co-workers.

What Back *et al.* do is to grow, on a tungsten substrate, single-domain films of iron that consist of one complete atomic layer of iron followed by a two-dimensional network of irregular patches. They then measure the critical exponents characterizing the approximately two-dimensional iron, and use these exponents to form the scaling equation of state. A remarkable feature of their work is that the system they study does not perfectly mirror the conditions of the Ising model, yet the measured scaling equation of state conforms almost perfectly to the calculated result<sup>4</sup>. This is a consequence of universality. The experiments in question are essentially two-dimensional, and the component 'magnetic moments' have sufficient anisotropy in their interactions that when acting collectively they behave as one-dimensional classical spins. Hence the experimental system mirrors a two-dimensional Ising model.

At one time, it was imagined that the 'scale-free' case was relevant to only a fairly narrow slice of physical phenomena<sup>5</sup>. However, the range of systems that apparently display power-law correlations has increased dramatically in recent years, ranging from base-pair correlations in DNA<sup>6</sup>, lung inflation<sup>7</sup> and interbeat intervals of the human heart<sup>7</sup>, to complex systems involving large numbers of interacting subunits that display 'free will', such as city growth<sup>8</sup> and even economics<sup>9</sup>. The principle of universality seems to be reflected in the empirical fact that these quite different systems can have remarkably similar critical exponents — perhaps because the 'interaction paths' between the constituent subunits in such extremely complex systems dominate the observed cooperative behaviour more than the detailed properties of the subunits themselves, just as Back *et al.* find that the simple two-dimensional Ising model describes well the interaction paths in their more complex real-life experiment. □

H. Eugene Stanley is in the Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, Massachusetts 02215, USA.

1. Stanley, H. E. *Introduction to Phase Transitions and Critical Phenomena* (Oxford Univ. Press, 1971).
2. Back, C. H. *et al.* *Nature* **378**, 597–600 (1995).
3. Stanley, H. E. & Ostrowsky, N. (eds) *Correlations & Connectivity: Geometric Aspects of Physics, Chemistry & Biology* (Kluwer, Dordrecht, 1990).
4. Gaunt, D. S. & Domb, C. *J. Phys. C* **3**, 1442–1461 (1970).
5. Peng, C.-K. *et al.* *Nature* **356**, 168–171 (1992).
6. Suki, B. *et al.* *Nature* **368**, 615–618 (1994).
7. Peng, C.-K. *et al.* *Chaos* **5**, 82–87 (1995).
8. Makse, H., Havlin, S. & Stanley, H. E. *Nature* **377**, 608–612 (1995).
9. Mantegna, R. N. & Stanley, H. E. *Nature* **376**, 46–49 (1995).



# A nice ring to the centrosome

Berl R. Oakley

ONE of the central mysteries in cytoskeletal research, the identity of the components of microtubule-organizing centres that nucleate microtubule assembly, seems to have been solved at last. Two remarkable papers on pages 578 and 638 of this issue<sup>1,2</sup> leave little doubt that ring-shaped complexes containing  $\gamma$ -tubulin are, in fact, the microtubule nucleators.

Microtubule-organizing centres (or MTOCs) play a central role in the regulation of microtubule assembly, and thus in all the functions carried out by microtubules (such as providing the framework for movement of chromosomes and organelles within the cell). Perhaps the best known MTOCs are the centrosomes of animal cells but analogous, though structurally different, MTOCs are found in a variety of phyla. MTOCs nucleate microtubule assembly and in doing so help to control the positions of microtubules in cells. Their ability to nucleate microtubule assembly is regulated in a cell-cycle-specific fashion, which helps coordinate the assembly of structures such as the mitotic apparatus with the cell cycle. MTOCs also serve two other functions. They establish the number of protofilaments in the microtubule at 13 and they establish the polarity of microtubules with the minus (slow growing) ends at the MTOC and the plus (fast growing) ends distal. All in all, identifying the components of MTOCs that nucleate microtubule assembly and establish microtubule polarity is fundamental to our understanding of mitosis, meiosis and the microtubule cytoskeleton. But it has been a difficult task.

A key step was the discovery of  $\gamma$ -tubulin<sup>3</sup>, which is distinct from, but related to, the major microtubule proteins,  $\alpha$ - and  $\beta$ -tubulin. It appears to be ubiquitous in eukaryotes<sup>4-6</sup>, and in fungal and animal cells it is located most obviously at MTOCs<sup>4-9</sup>. Experiments involving the disruption of  $\gamma$ -tubulin genes<sup>7,8</sup>, and blockage<sup>9,10</sup> and depletion<sup>11</sup> experiments using antibodies against  $\gamma$ -tubulin, all indicated that  $\gamma$ -tubulin is essential for microtubule assembly from MTOCs. These results, along with the genetic experiments that led to the discovery of  $\gamma$ -tubulin, suggested that rings of  $\gamma$ -tubulin nucleate microtubule assembly<sup>6,7</sup>. There was no direct evidence, however, that such rings existed, though earlier this year it was shown that *in vitro*-translated  $\gamma$ -tubulin binds to the minus ends of microtubules with a stoichiometry consistent with this model<sup>12</sup>.

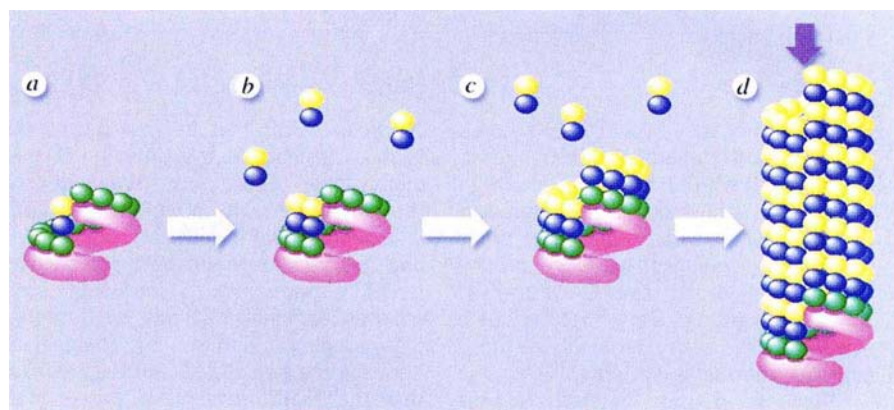
In a technical *tour de force* employing immunoelectron microscopic tomography, Moritz *et al.*<sup>2</sup> have now demonstrated that  $\gamma$ -tubulin is, indeed, located in ring structures within the pericentriolar material of

isolated centrosomes. When microtubules are assembled from these centrosomes, moreover,  $\gamma$ -tubulin complexes are located at their minus ends, just as expected if the rings nucleate microtubule assembly.

In a complementary and equally impressive study, Zheng *et al.*<sup>1</sup> have built upon the finding that, in unfertilized *Xenopus* eggs,  $\gamma$ -tubulin exists in 25S complexes<sup>10</sup>. As *Xenopus* eggs are loaded with components necessary for repeated rounds of mitosis, it was possible that these were functional  $\gamma$ -tubulin complexes ready to be incorporated into centrosomes after fertilization. In spite of the

contains a break or offset. In recent years it has become apparent that microtubules assemble in solution to form what is called a B-lattice structure. Although the exact nature of the microtubule lattice may seem esoteric, it has important consequences. The end of a B-lattice microtubule cannot be flat. A flat ring structure could not easily serve as a template for microtubule assembly, but a ring with an offset would do very nicely. Zheng *et al.* also find that each  $\gamma$ -tubulin ring complex contains approximately one molecule each of  $\alpha$ - and  $\beta$ -tubulin. They propose that an  $\alpha/\beta$  tubulin dimer is present at the offset in the  $\gamma$ -tubulin ring. This would provide stabilizing lateral interactions for a tubulin dimer binding to this region, and so facilitate the nucleation of microtubule assembly.

Two other features of this model (and



A model for microtubule nucleation by the  $\gamma$ -tubulin ring complex (modified from Zheng *et al.*<sup>1</sup>). The  $\gamma$ -tubulin ring complex (a) contains 13  $\gamma$ -tubulin molecules (green), one  $\alpha/\beta$  tubulin dimer (yellow and blue) at the offset in the ring, and several other proteins (pink) identified so far only as bands on gels. Initial binding of a tubulin dimer adjacent to the ring complex tubulin dimer (b) is stabilized by bonds to  $\gamma$ -tubulin and to the ring complex dimer. Additional dimers bind sequentially (c) forming a ring of dimers on the  $\gamma$ -tubulin ring complex. Binding of more tubulin dimers creates a complete B-lattice microtubule (d). Here the tubulin dimers are arranged to form a B-lattice, the most likely structure for the microtubule. In the B-lattice,  $\alpha$ -tubulin molecules are beside other  $\alpha$ -tubulin molecules and  $\beta$ -tubulin molecules are beside other  $\beta$ -tubulin molecules except at a seam (arrow). Such  $\gamma$ -tubulin ring complexes would serve as a perfect template to nucleate the assembly of a B-lattice microtubule. In addition, if  $\gamma$ -tubulin binds to  $\alpha$ - or  $\beta$ -tubulin but not both molecules, the  $\gamma$ -tubulin ring complexes would establish the polarity of microtubules. If, for example,  $\gamma$ -tubulin binds to  $\beta$ -tubulin, the  $\beta$ -tubulin monomer of each  $\alpha/\beta$  tubulin dimer would always lie towards the microtubule-organizing centre. (Figure prepared by Tetsuya Horio, Kazusa DNA Institute, Kisarazu, Japan.)

low abundance of the complexes, Zheng *et al.* have developed a procedure, including a critical antibody-affinity step, that allows purification of the complexes in an active form. The complexes have a ring structure and contain about 10–13  $\gamma$ -tubulin molecules as well as several other proteins. Most significantly, these rings nucleate microtubule assembly and cap the minus ends of microtubules *in vitro*. In combination, these studies should convince even the most sceptical that  $\gamma$ -tubulin ring complexes are the long-sought minus-end nucleators of microtubule assembly.

Zheng *et al.* suggest an attractive modification of the original view of  $\gamma$ -tubulin nucleation of microtubule assembly (reproduced here in slightly modified form). The purified  $\gamma$ -tubulin ring complex

the original on which it is based) are very attractive. One is that the presence of 13  $\gamma$ -tubulin molecules in the ring complex would establish the protofilament number of microtubules. The second is that, as the figure shows, if  $\gamma$ -tubulin binds to  $\alpha$ - or  $\beta$ -tubulin, but not both, the complexes would establish the polarity of the microtubules that extend from them.

These findings allow us to look at the regulation of microtubule assembly in a new light. The spatial location of microtubule assembly is almost certainly controlled, in part, by the positioning of  $\gamma$ -tubulin ring complexes, but what controls the positioning of these complexes remains unknown. The assumption is that  $\gamma$ -tubulin ring complexes bind to specific proteins at centrosomes and other

MTOCs, but the challenge now is to identify these proteins and find out how their positions are established.

Part of the temporal regulation of microtubule assembly could involve translocation of  $\gamma$ -tubulin to and from centrosomes<sup>4,5,13</sup>, but it is clear that, at least in fungi,  $\gamma$ -tubulin is present at MTOCs when no spindle is present<sup>7,8</sup>. The microtubule nucleating ability of the  $\gamma$ -tubulin ring complexes thus appears to be turned on and off by the cell-cycle regula-

tory machinery. The most obvious model is that the microtubule-nucleating ability might be turned off in interphase and, at the G2 to M transition, a kinase involved in cell-cycle regulation (for instance the p34<sup>cdc2</sup> kinase) might phosphorylate  $\gamma$ -tubulin, or some other proteins in the nucleating complex, and turn microtubule-nucleating ability on.

The results of Zheng *et al.* point to another possibility, however. Although the  $\gamma$ -tubulin ring complexes are present in the

cytoplasm of unfertilized *Xenopus* eggs, they do not nucleate microtubule assembly until they become part of centrosomes, after fertilization. So it is somewhat surprising that ring complexes purified from unfertilized eggs nucleate microtubule assembly. One highly speculative, but intriguing, possibility is that in the egg they are complexed with an inhibitor molecule that blocks microtubule nucleation. This molecule might be removed artificially during the purification process. In a more natural situation, phosphorylation of such an inhibitor molecule at the onset of mitosis might cause it to release from the  $\gamma$ -tubulin ring complex and turn on the nucleating capacity of the complex. □

Berl R. Oakley is in the Department of Molecular Genetics, Ohio State University, Columbus, Ohio 43210, USA.

1. Zheng, Y., Wong, M. L., Alberts, B. & Mitchison, T. *Nature* **378**, 578–583 (1995).
2. Moritz, M., Braunfeld, M. B., Sedat, J. W., Alberts, B. & Agard, D. A. *Nature* **378**, 638–640 (1995).
3. Oakley, C. E. & Oakley, B. R. *Nature* **338**, 662–664 (1989).
4. Zheng, Y., Jung, M. K. & Oakley, B. R. *Cell* **65**, 817–823 (1991).
5. Stearns, T., Evans, L. & Kirschner, M. *Cell* **65**, 825–836 (1991).
6. Oakley, B. R. *Trends Cell Biol.* **2**, 1–5 (1992).
7. Oakley, B. R., Oakley, C. E., Yoon, Y. & Jung, M. K. *Cell* **61**, 1289–1301 (1990).
8. Horio, T. *et al.* *J. Cell Sci.* **99**, 693–700 (1991).
9. Joshi, H. C., Palacios, M. J., McNamara, L. & Cleveland, D. W. *Nature* **356**, 80–83 (1992).
10. Stearns, T. & Kirschner, M. *Cell* **76**, 623–637 (1994).
11. Félix, M.-A., Antony, C., Wright, M. & Maro, B. *J. Cell Biol.* **124**, 19–31 (1994).
12. Li, Q. & Joshi, H. C. *J. Cell Biol.* **311**, 207–214 (1995).
13. Lajoie-Mazenc, I. *et al.* *J. Cell Sci.* **107**, 2825–2837 (1994).

## OBITUARY

## Jeffries Wyman (1901–95)

JEFFRIES Wyman, who died at his home in Paris on 4 November 1995, was a biophysicist who contributed much to our understanding of haemoglobin and its interactions with oxygen and other ligands, as a result of his work at Harvard and later at the University of Rome. His writings show that he was a master of the thermodynamics of these complex interacting systems.

Following in the footsteps of his grandfather, the first Jeffries Wyman, who was one of the founding members of the National Academy of Sciences, Jeffries started his research career at Harvard. He first looked into the dielectric constants of aqueous solutions of amino acids and peptides, which were known to be highly polar molecules. The dielectric constants of their aqueous solutions proved to be much greater than that of water, and the values were systematically related to structure. Jeffries brought his findings and interpretations to Lars Onsager, which stimulated Onsager to produce the first satisfactory theory of highly polar liquids, later improved by John G. Kirkwood.

In 1937 his first haemoglobin paper appeared, with Bernard German as coauthor, on the pH titration curves of oxy- and deoxyhaemoglobin. Measurements on oxyhaemoglobin were now possible because glass electrodes had only recently become available. Jeffries made a key thermodynamic analysis of the data, a forerunner for his later and far more extensive treatments. One such analysis, of linked functions and reciprocal relations, appeared in 1948, in the context of a review of haem proteins, and proved a powerful tool for interpreting experimental data.

It was while visiting fellow scientists in

Japan in 1950 that he had a sudden flash of insight. At the time, no three-dimensional data were available on the structure of any protein, but it had long been known that the crystals of oxy- and deoxyhaemoglobin belonged to different classes. Felix Haurowitz had recently demonstrated this by showing the breakup during the transition between the two states. Jeffries realized that the rearrangement of the crystal structures implied a change, on ligand binding or release, in the conformation of the molecules themselves. As he was to put it in the published paper “...the reason why certain acid groups are affected by oxygenation is simply the alteration in their position and environment which results from the change of configuration of the hemoglobin molecule as a whole accompanying oxygenation”.

This was a clear statement of the concept that Jacques Monod was later to christen allostery. Jeffries developed the idea at Harvard with his student David Allen, and the paper was eventually published in 1951. Today, in the wake of overwhelming and detailed confirmation by the work of Max Perutz, the idea seems so obvious that we take it for granted.

Jeffries resigned from Harvard a year later to become the first Science Attaché of the American Embassy in Paris. After three years he went on to serve as Director of the UNESCO Science Cooperation Office for the Middle East, with headquarters in Cairo, with responsibilities extending from Morocco to Pakistan.

When that assignment came to an end in 1960, he was for a time uncertain what to do. His doubts were resolved by a visit from a brilliant and

enthusiastic young Italian, Eraldo Antonini, who urged Jeffries to join his group at the University of Rome, working on haemoglobin on problems closely related to those studied by Jeffries at Harvard. He agreed to come for one year and stayed for twenty-four, in what proved to be the most productive years of his research life.

A visit to Paris in 1964 to discuss allostery with Monod, whom he knew from his years at the American Embassy, and J.-P. Changeux, led to the paper by Monod, Wyman and Changeux published in 1965 on a model for allosteric enzymes, including haemoglobin as an “honorary enzyme”. This was one of the most influential of the many models proposed to account for such phenomena. Jeffries followed this by a large and rather formidable paper on “allosteric linkage” in 1967, and continued to work on allostery for the rest of his life.

He was coauthor of two books: *Biophysical Chemistry* (1958), which he and I wrote together, dealing with topics we had presented at Harvard; and *Binding and Linkage: Functional Chemistry of Biological Macromolecules* (1990) with Stanley J. Gill. The latter was the product of some twelve years of thought and rewriting, in constant consultation with critical friends. It represents a final legacy of the two authors, both of whom have now died, and a powerful tool for investigators in this important and difficult field.

John T. Edsall

John T. Edsall is in the Department of Molecular and Cellular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138-2092, USA.

# Mapping to the point

Richard Korn and Yashin Ramkissoon

FOR all that modern molecular genetics makes identification of disease genes easier, nothing can substitute for detailed clinical observation, careful selection of individuals for study, inspiration and hard work. These virtues are amply illustrated in a paper by Ellis *et al.*, just published in *Cell*<sup>1</sup>, which describes the positional cloning of the gene responsible for Bloom's syndrome (BS)<sup>2</sup>; this is a rare, autosomal recessive condition, characterized by short stature, immunodeficiency, photosensitive skin lesions and an increased risk of developing all types of cancer.

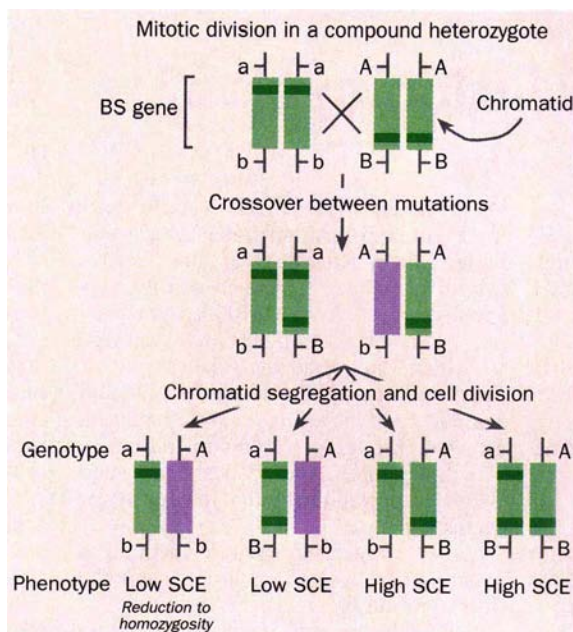
The work has its foundation in over 30 years of detailed cytological and biochemical observations of people with BS. During this time German<sup>3</sup> compiled a worldwide registry of affected individuals (currently 179 are known), and produced and studied immortalized cell lines from many of them. The consequent accumulation of knowledge and resources, together with modern molecular techniques (and some sharp thinking), allowed the gene to be rapidly identified by means of a strategy unique to the disease.

Standard linkage mapping has not been feasible with BS because so few people are affected. Ellis, German and colleagues had previously used a method known as homozygosity mapping to define the BS locus<sup>4</sup> (*BLM*, wild-type allele; *blm*, mutant allele), affected children of consanguineous parents being expected to be homozygous by descent at and around *BLM* significantly more often than elsewhere in the genome. One marker mapping within the *FES* proto-oncogene, at position 26.1 on the long arm of chromosome 15 (15q), was found to be homozygous in 25/26 patients examined. Subsequent analysis revealed allelic linkage disequilibrium<sup>5</sup> between the markers *FES*, D15S127 and *BLM*, placing *BLM* close to *FES*.

One of the characteristics of BS is a high level of genomic instability, marked by chromosome breaks, gaps, translocations, a high spontaneous mutation rate and a high frequency of both homologous chromatid and sister-chromatid exchanges (SCE), all of which are consistent with a defect in DNA replication. Cells from patients with BS have about ten times the number of SCEs compared with normal cells<sup>6</sup>, a feature that is diagnostic of the disease. But one perplexing observation

has been that some BS patients have a minor population of lymphocytes with normal SCE levels, and none with intermediate levels. When fused with high-SCE cells, the low-SCE cells restore a normal phenotype in the same way that fusion to a normal cell does. But how could these apparently normal, low-SCE cells arise?

A reasonable assumption is that such cells are the result of some sort of somatic mutation event. Many mechanisms have



Intragenic somatic recombination in Bloom's syndrome (BS). The high frequency of recombination can generate a wild-type *BLM* allele (purple) in compound heterozygotes, accounting for the minor population of lymphocytes with normal levels of sister chromatid exchange (SCE) seen in these individuals. A/a and B/b are polymorphic microsatellite markers respectively located proximal and distal to the *BLM* locus. If intragenic recombination is occurring, reduction to homozygosity of marker b distal to *BLM* should occur in 50 per cent of low-SCE cells. This phenomenon was used by Ellis *et al.*<sup>1</sup> to delineate the *BLM* critical region. Mutations, dark green; *blm* mutant alleles, light green.

been postulated<sup>7</sup> but none has found support until recently. As the mixed high-SCE/low-SCE phenotype occurs almost exclusively in offspring of non-consanguineous parents, Ellis *et al.* reasoned that these patients must be compound heterozygotes—that is, have different mutant alleles of *blm* inherited from each parent. They subsequently proposed a model of intragenic somatic recombination<sup>8</sup> to explain the revertant phenotype in compound heterozygotes: recombination occurs *within* the BS gene, transferring both mutations to the same chromatid and simultaneously creating a wild-type allele on the other chromatid (see figure).

One expectation of this hypothesis is

loss of heterozygosity (or reduction to homozygosity) on chromosome 15 distal to *BLM* in half of the low-SCE recombinants. Low-SCE cell lines from 5 of 11 individuals indeed showed reduction to homozygosity of all markers distal to 15q26.1, and remained heterozygous proximal to 15q26.1. Cytogenetically there was no loss of distal chromosome 15q. This is strong evidence that intragenic somatic recombination accounts for the revertant phenotype.

Recognizing that the high rate of mitotic recombination could provide an approach for cloning of *BLM*, Ellis *et al.*<sup>1</sup> delineated a minimal region of 15q containing the gene by typing low-SCE cell

lines with polymorphic markers, and looked for the change from heterozygosity to homozygosity proximal to *BLM*. The resolution of mapping was limited only by the density of available markers in this region. Using this method, which they term somatic crossover-point mapping, the critical region was narrowed down to 250 kilobases, from which a full-length complementary DNA was isolated. Analysis of the open reading frame in patients with BS revealed a range of nonsense, missense and frame-shift mutations, showing that here at last was the BS gene.

The predicted *BLM* protein has homology to the seven helicase domains of the *Escherichia coli* recQ and human REQL proteins<sup>1</sup>, and it is likely that it will share their DNA-dependent ATPase and DNA helicase activities. Loss of *BLM* function causes a variety of genomic and biochemical aberrations. It is possible that loss of helicase activity generates abnormal DNA structures during replication which indirectly affect the activity of other DNA-binding proteins or activates repair mechanisms (the latter potentially accounting for the high rates of SCE). Alternatively, direct interactions of *BLM* with other proteins may be required for their correct function. Interestingly, genes associated with other genomic instability syndromes, for instance those involved in Cockayne's syn-

1. Ellis, N. *et al.* *Cell* **83**, 655–666 (1995).
2. McKusick, V. *Mendelian Inheritance in Man*, 10th Edn (Johns Hopkins Univ. Press, Baltimore, 1992).
3. German, J. & Passarge, E. *Clin. Genet.* **35**, 57–69 (1989).
4. German, J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **91**, 6669–6673 (1994).
5. Ellis, N. *et al.* *Am. J. hum. Genet.* **55**, 453–460 (1994).
6. German, J. *et al.* *Am. J. hum. Genet.* **29**, 248–255 (1977).
7. Weksberg, R. *et al.* *Am. J. hum. Genet.* **42**, 816–824 (1988).
8. Ellis, N. *et al.* *Am. J. hum. Genet.* **57**, 1019–1027 (1995).
9. Troelstra, C. *et al.* *Cell* **71**, 939–951 (1992).
10. Marx, J. *Science* **266**, 728–730 (1994).
11. Goodfellow, P. N. *Nature* **377**, 285–286 (1995).



drome group B (ERCC6)<sup>9</sup> and two of the seven complementation groups of xeroderma pigmentosum (XPB and XPD)<sup>10</sup>, encode helicases that take part in nucleotide excision repair.

The success of the human genome project has led to rapidly expanding resources for the gene mapping community, and of these the EST (expressed sequence tag) database is becoming increasingly important. The mapping of ESTs to their position in the genome to produce a transcript map will have enormous impact on the positional cloning of genes<sup>11</sup>. In the foreseeable future once a gene has been mapped to within 1–2 centimorgan (or megabases), details of the 50 or so cDNAs

which map to that region will be available from a database providing both positional and functional information. All candidate genes could then be systematically screened for mutations. Although this 'factory' approach will be both scientifically profitable and hugely successful, there are those who mourn the passing of intelligent and innovative gene cloning, as demonstrated in the identification of *BLM*. □

*Richard Korn and Yashin Ramkissoon are in the Laboratory of Human Molecular Genetics, Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK.*

## INTERPLANETARY DUST

# A driver of glaciation cycles?

Donald E. Brownlee

As the Earth orbits the Sun, it collides with vast numbers of small particles amounting to over 10,000 tons of accreted extraterrestrial material a year. The bulk of this mass is composed of fragments of comets and asteroids, between 10 micrometres and 1 millimetre in diameter. This rain of primitive material from the outer Solar System onto the Earth may have played a role in the development of life<sup>1</sup>, and evidence presented on page 600 of this issue suggests that it may also play a role in terrestrial glaciation cycles. Here, Farley and Patterson<sup>2</sup> use <sup>3</sup>He abundances measured in a deep-sea sediment core to monitor variations in the accretion rate of interplanetary dust particles (IDP). What they have found are temporal changes that correlate with the 100-kyr periodicity in the succession of ice ages on the Earth, one of the so-called 'Milankovitch' cycles.

To see why this finding will make climate scientists sit up, a little background on the glacial-interglacial cycles is needed. The basic idea developed by Milankovitch was that cyclical changes in the Earth's orbital eccentricity, in the tilt of its axis (obliquity) and in its precession, affect the distribution of solar radiation and thus draw in their wake changes to the growth and decay of ice sheets. At least some of the dominant frequencies seen in the geological record of ice ages do seem to be plausibly accounted for in this way, the periodicities at 19 and 23 kyr being attributable to precession and the 41-kyr periodicity to obliquity changes. But in the past 800 kyr, the strongest signal in the climate record has been at a frequency of 100 kyr, and although this in principle corresponds with the eccentricity cycle, variations in eccentricity should produce the weakest of the orbital effects on incident solar radiation.

Earlier this year, Muller and MacDon-

ald proposed<sup>3</sup> that the answer to this puzzle might lie in a hitherto neglected parameter: the inclination of the Earth's orbital plane. Like the eccentricity, this should vary with roughly 100-kyr periodicity, but the changes would not be expected to affect the incident solar radiation directly. Instead, Muller and MacDonald suggested that the climate cycle might be affected through variations in the accretion rate of interplanetary dust. Farley and Patterson have now found evidence for such a variation.

Dust in the interplanetary medium is generated by asteroid collisions and the disintegration of comets. On timescales of 10<sup>4</sup> to 10<sup>6</sup> years, individual submillimetre particles are lost from the inner Solar System by collisional disruption and orbital decay induced by radiation pressure (the Poynting–Robertson effect). On average, the production and destruction mechanisms must balance, and one might expect that the particle flux reaching the Earth would be in a quasi-steady state with superimposed fluctuations due to particle generation episodes such as the appearance of large, active comets. Additional complications include the formation of a dust belt just outside the Earth's orbit<sup>4</sup> and variations of the Earth's orbital parameters.

As suggested by Muller and MacDonald<sup>3</sup>, changes in orbital inclination could modulate the flux of accreted dust with a 100-kyr period. But the new results measured by Farley and Patterson indicate that the accretion rate is lowest at times of highest inclination; in other words, the phase of variation is the opposite of what would be expected. The effect, if confirmed, cannot merely be a simple modulation caused by the inclination sweeping through a distribution of particles concentrated to the ecliptic.

IDPs contain substantial amounts of

helium (up to concentrations equivalent to 1 cm<sup>3</sup> per gram at STP). As was discovered over three decades ago<sup>5</sup>, the isotopic composition of this helium is very different from that of most terrestrial helium found in deep-sea sediments. The isotopic composition of the IDP He is similar to the primordial ratio whereas the terrestrial component is largely radiogenic and contains mainly <sup>4</sup>He. The lighter isotope <sup>3</sup>He is an excellent tracer for extraterrestrial matter and is retained in sediments for long periods of time<sup>6</sup>, so Farley and Patterson use <sup>3</sup>He abundances throughout a sediment core to trace the accretion of extraterrestrial material.

A most interesting aspect of the data is that the implied flux is high and has a 100-kyr variability only during the Quaternary period, which began around 2 million years ago. The flux was much lower for all of the preceding 70 million years. If correct, this indicates that the recent 100-kyr He variability correlates with the recent development of the 100-kyr climate cycles. It is odd, however, that the highest He abundances are found at the top of the sediment column and it is possible that the results may be influenced by partial He loss in older sediments.

Overall, these data are interesting and important. If a true causal link can be established between the IDP flux and any climate cycle, then the real question will be, how this can happen? How can meteoritic material have an effect when it constitutes only a minuscule fraction of the total particulate mass loading of the atmosphere? It is unlikely that small particles that the He measurements actually detect in sediments could have much effect. By virtue of their small size and moderate entry velocity, these survive atmospheric entry intact. But high-velocity particles and those larger than 0.1 micrometre in diameter are partly or totally vaporized in the mesosphere, and the resulting vapours and fine condensates<sup>7</sup> may play a more effective role. Although the timescales are different, these findings are reminiscent of a mechanism suggested by Bowen many years ago for linking rainfall with meteor showers<sup>8</sup>. Farley and Patterson's results remind us once again that our planet's climate is at the mercy of its wider environment. □

*Donald E. Brownlee is in the Department of Astronomy, University of Washington, Seattle, Washington 98195, USA.*

1. Anders, E. *Nature* **342**, 255–257 (1989).
2. Farley, K. A. & Patterson, D. B. *Nature* **378**, 600–603 (1995).
3. Muller, R. A. & MacDonald, G. J. *Nature* **377**, 107–108 (1995).
4. Dermott, S. F. *et al.* *Nature* **369**, 719–723 (1994).
5. Merrihue, C. *Ann. N. Y. Acad. Sci.* **119**, 351–367 (1964).
6. Farley, K. A. *Nature* **376**, 153–156 (1995).
7. Huxten, D. M., Turco, R. P. & Toon, O. B. *J. Atmos. Sci.* **37**, 1342–1357 (1980).
8. Bowen, E. G. *J. Meteorol.* **13**, 142–151 (1956).

# From out of the blue

George Sugihara

NEARLY 2,000 years after accounts in the Book of Exodus of the locust plagues that God brought down upon the Egyptians, Chinese bureaucrats of the T'ang Dynasty (618–907 BC) began organizing one of the most ambitious entomological data-collection projects in history. As annual reports from district officials have been faithfully recorded to the present day, a remarkable time series was constructed<sup>1</sup> for the degree of locust infestation in the form of a severity index (1–10 based on density and spatial extent, a sort of non-logarithmic Richter scale) (Fig. 1). The original aim, which has yet to be completely fulfilled, was to use this information to help forecast locust outbreaks.

Little did the T'ang officials realize that, 1,300 years later, these data as well as their country's politics would be found

giving rise to an apparent paradox: observations for natural populations, such as those in refs 2 and 3, are redshifted, whereas theoretical models produce blue spectra.

This paradox has its roots in the debate between the so-called 'climate' and 'biotic' schools of population regulation that has polarized thinking among ecologists since the early part of this century. Although many economic entomologists favoured the view that external climatic factors dominated populations, others (especially those studying longer-lived and larger creatures such as birds and mammals) held that internal biotic mechanisms such as competition and predation are paramount. This debate has been reborn in various guises: density-independence versus density-dependence; nonequilibrium versus equilibrium; null models versus

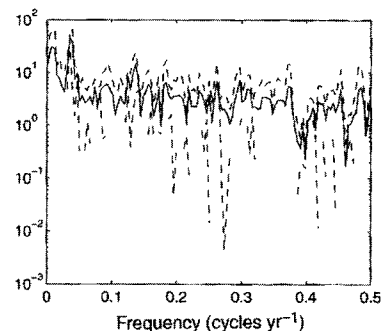


FIG. 2 The power spectrum for the migratory locust time series showing a shift towards the red. A moving window of 256 years was used to estimate the spectrum, and the error envelope (dashed line) corresponds to the 95% confidence interval.

wet periods, which promote nymphal survival. Moreover, climate variables themselves are often redshifted on this timescale (with interdecadal bumps).

So although redshifted observations align with the climate school, the blue-shifted output of classical autonomous population models align with the biotic school. Chaos in these simple models is achieved when self-regulation has run amuck (that is, when the time-delayed birth/death response is fast, causing the system to overshoot and undershoot a target equilibrium to an extreme that becomes aperiodic<sup>6</sup>).

Indeed, the blueshift that Cohen observes should tend to be true of one-dimensional, unimodal maps in general, where there is anticorrelation between values at successive time steps. Thus, low values tend to be followed by highs and vice versa, giving rise to short-term anti-correlation or maximal power at the Nyquist frequency. In the chaotic regime, these correlations will fall off with the Lyapunov horizon, giving rise to an apparent blueshift. To achieve the intermittency that would push the spectra for these models towards red, they would need to be sensitively tuned to specific pathological values. Chaos with red spectra can arise more readily in higher dimensional models. So the full story behind the ragged ups and downs that many populations in nature exhibit is not likely to be explained as examples of simple chaos from classical difference equation models (that is, stationary, unimodal single-species maps). ▶

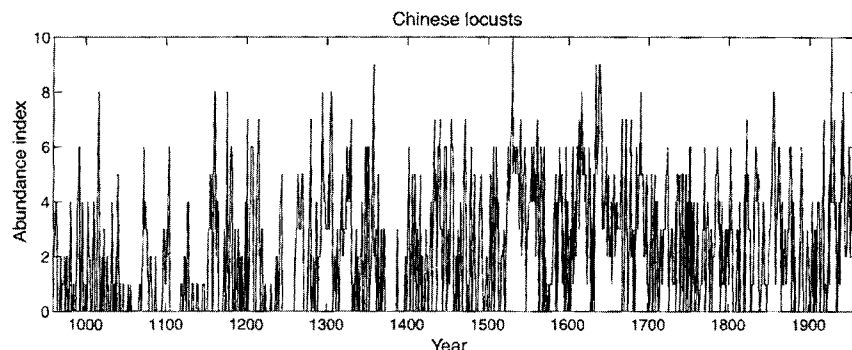


FIG. 1 A 1,000-year record (957–1956) of locust (*Locusta migratoria*) abundance in China, reproduced from ref. 1. This is an excerpt of the original 1,300-year total abundance record for all census regions combined.

to be shifted to the red (Fig. 2). That is to say, in common with many other population data on wild stocks, for instance those of Diamond and May<sup>2</sup>, and Pimm and Redfearn<sup>3</sup>, this ancient record of locust booms and busts possesses a redshifted power spectrum. This means that population variability appears to increase at the longer timescales. Thus, two population values a century apart will on average differ more than two population values a decade apart, and so on. And so what?

On page 610 of this issue<sup>4</sup>, Joel Cohen presents a new dilemma in population modelling that pits this observation against classical theory. As Cohen points out, the ubiquity of such a  $1/f$  signature, or shift in power to the low frequencies, is at odds with his simple analysis of prevailing autonomous population models (some of which are used in resource management). Cohen shows how these models tend to have blueshifted power spectra in the chaotic regime (more variation in the higher frequencies; adjacent generations tend to be more different than distant ones),

biogeographical pattern; physics of transport versus real population growth; noise versus chaos. It is a practical problem for resource managers on a daily basis. For example, whether or not commercial fishing can be sustained depends critically on whether specific fish populations are behaving in a predominantly density-dependent or density-independent manner<sup>5</sup>, and the form of the density dependence is especially crucial.

Insofar as the Chinese locust data measure the action of locusts, and not the sociological trends of census takers, the redshift reported here and the many cases of short-term redshifted population data surveyed elsewhere<sup>3</sup> seem to be in step with the views of the climate school (though as Diamond and May point out<sup>2</sup>, with short-term measurements it is difficult to discount other sources of 'nonstationarity', such as human disturbance, as the cause of the redshift). Indeed the only published analysis<sup>1</sup> of the locust data suggests a significant correlation between outbreaks and climatic factors, especially

1. Ma, S. C. *Acta Ent. Sinica* **8**, 1–40 (1958).
2. Diamond, J. M. & May, R. M. *Science* **197**, 266–270 (1977).
3. Pimm, S. L. & Redfearn, A. *Nature* **334**, 613–614 (1988).
4. Cohen, J. E. *Nature* **378**, 610–612 (1995).
5. Ludwig, D., Hilborn, R. & Walters, C. *Science* **260**, 17 (1993).
6. May, R. M. *Nature* **261**, 459–467 (1976).
7. Tong, H. *Nonlinear Time Series Analysis: A Dynamical Systems Approach* (Oxford Univ. Press, 1990).
8. Rand, D. A. & Wilson, H. B. *Proc. R. Soc. B* **246**, 179–184 (1991).
9. Sugihara, G. *Phil. Trans. R. Soc. B* **348**, 477–495 (1994).
10. Steele, J. H. *Nature* **313**, 355–358 (1985).

This leaves us with at least three alternatives: first, that natural population fluctuations are not chaotic and/or are nonstationary; second, that the models are fundamentally flawed; or, third, that environmental forcing needs to be incorporated. Such forcing can easily push a stable map into the chaotic regime<sup>7-9</sup>, and Steele<sup>10</sup> has suggested that environmental forcing (red noise) superimposed on models containing multiple stable states may

lead to the kind of intermittency that could produce a redshift. Whatever the case, Cohen's article poses a dilemma that should stimulate some lively discussion around what remains one of the classic unresolved issues in ecology. □

George Sugihara is at the Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0202, USA.

## PLANT BIOLOGY

## At the roots of nutrition

Julian I. Schroeder

PLANTS grow and develop while remaining rooted in one place. Therefore, identification of the molecular mechanisms through which plants accumulate the often sparse soil macronutrients essential for growth has been a central aim in plant biology. Two papers in this issue add further to our knowledge of such mechanisms. Harrison and van Buuren (page 626<sup>1</sup>) report the first cloning of a mycorrhizal phosphate-

(patch) clamp and molecular cloning studies having provided first insights into the biophysical uptake mechanisms and molecular structures of key plant transporters for the essential 'NPK' macronutrients (nitrogen, phosphorus and potassium). These include low-affinity<sup>3</sup> and high-affinity nitrate-uptake transporters<sup>4</sup> and high-affinity ammonium transporters<sup>5</sup>. Furthermore, low-affinity K<sup>+</sup> uptake channels<sup>6-8</sup> and high-affinity K<sup>+</sup> transporters<sup>9,10</sup> have been characterized in cell biological, biophysical and molecular terms. But the molecular mechanisms for phosphate uptake have remained unknown.

Harrison and van Buuren<sup>1</sup> describe the isolation of a complementary DNA encoding a high-affinity phosphate-uptake transporter from vesicular arbuscular (VA) mycorrhizal fungi. The family of VA mycorrhizal fungi form symbiotic associations with most higher plants, including agricultural crops. This association greatly increases plant phosphate uptake (see figure), growth and crop yields<sup>11</sup>, the fungus getting carbon in return. The phosphate transporter, from VA *Glomus versiforme* mycorrhizae, named GvPT, was cloned by homology to the PHO84 phosphate transporter from yeast<sup>12</sup>. The GvPT sequence encodes a membrane protein, which complements a yeast mutant defective in high-affinity phosphate uptake. GvPT also has a high affinity for phosphate ( $K_m$  value of about 18  $\mu$ M) required for nutrient accumulation. The transporter appears to be expressed predominantly in the external hyphae of the mycorrhizae; VA fungi have extended hyphae that can reach into undepleted soil regions and are crucial for transfer of phosphate and other minerals from the soil to the host plants<sup>11</sup>.

We now have fresh challenges. For example, further research will be needed to characterize the mechanisms of phosphate transfer from mycorrhizae to roots, to identify plant (non-fungal) phosphate uptake transporters, and to determine the effects of transgenic manipulation of GvPT on phosphate nutrition. On the practical side, further characterization

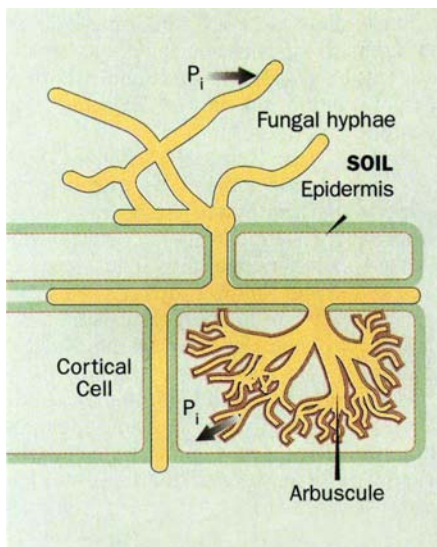
may allow investigation of the use of GvPT/PHO84 homologues for biological removal of leached phosphate.

Nitrogen is a macronutrient that often limits plant growth and crop yields. Leguminous plants have built-in symbiotic factories, root nodules, which contain nitrogen-fixing bacteroids inside organelles (symbiosomes) of the root cells. The bacteroids produce excess ammonia (NH<sub>3</sub>) which accumulates in the symbiosome space. Symbiosomes are surrounded by a specialized organellar membrane, the peribacteroid membrane, but the mechanisms for NH<sub>3</sub> or ammonium ion (NH<sub>4</sub><sup>+</sup>) uptake across this membrane into the root cytosol have remained unknown.

Tyerman *et al.*<sup>2</sup> address this question in a first patch-clamp analysis of the peribacteroid membrane. They report a prominent, non-selective monovalent cation current in the peribacteroid membrane that is activated by negative potentials on the plant cytosol side of the membrane. As the transporters are most permeable to ammonium, the negative potential would cause NH<sub>4</sub><sup>+</sup> uptake into plants. The cation transporter has a low affinity for NH<sub>4</sub><sup>+</sup> uptake ( $K_m$  of about 37.5 mM), which correlates with estimated NH<sub>4</sub><sup>+</sup> levels of about 20 mM in the symbiosome space.

Reversal potentials for the NH<sub>4</sub><sup>+</sup> uptake currents shifted in a strictly Nernstian manner at a symbiosome NH<sub>4</sub><sup>+</sup> concentration of between 3 and 150 mM. This shows that the ammonium uptake pathway is passive, which is the thermodynamic definition of a channel. The passive ammonium uptake channels differ from the plant high-affinity H<sup>+</sup>-coupled NH<sub>4</sub><sup>+</sup> transporters, which show a  $K_m$  of about 10  $\mu$ M (ref. 5).

To approximate the conductance and density of the underlying cation channels, the authors have pursued fluctuation analyses (which can be compared to analysing the noise of uniform raindrops falling on a tent roof to estimate the size and density of drops). This analysis showed that the ammonium channels have unusual, 'non-Lorentzian' properties, pointing to the existence of multiple states, cooperativity among channels or



Symbiosis in plant nutrition. A vesicular arbuscular mycorrhizal fungus invades root cortical cells and facilitates phosphate (P<sub>i</sub>) uptake from soil.

uptake transporter proposed to function in plant phosphate nutrition; and Tyerman *et al.* (page 629<sup>2</sup>) describe the identification of a mechanism for nitrogen accumulation from nitrogen-fixing root nodules.

Macronutrients are usually added to fertilizers, because their limited availability in soils is responsible for reduced plant growth. But some consequent environmental problems, such as large energy use for fertilizer production and groundwater contamination caused by leaching of fertilizers, have added impetus to the quest to identify associated nutrient transporters. A great deal of progress has been made, voltage

1. Harrison, M. J. & van Buuren, M. L. *Nature* **378**, 626-629 (1995).
2. Tyerman, S. D., Whitehead, L. F. & Day, D. A. *Nature* **378**, 629-632 (1995).
3. Tsay, Y. F., Schroeder, J. I., Feldmann, K. A. & Crawford, N. M. *Cell* **72**, 705-713 (1993).
4. Quesada, A., Galvan, A. & Fernandez, E. *Plant J.* **5**, 407-419 (1994).
5. Ninnemann, O., Jauniaux, J. C. & Frommer, W. B. *EMBO J.* **13**, 3464-3471 (1994).
6. Sentenac, H. *et al. Science* **256**, 663-665 (1992).
7. Anderson, J. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 3736-3740 (1992).
8. Schachtman, D. P. *et al. Science* **258**, 1654-1658 (1992).
9. Schachtman, D. P. & Schroeder, J. I. *Nature* **370**, 655-658 (1994).
10. Maathuis, F. J. M. & Sanders, D. *Proc. natn. Acad. Sci. U.S.A.* **91**, 9272-9276 (1994).
11. Smith, S. E. & Gianinazzi-Pearson, V. A. *Rev. Pl. Physiol. Pl. molec. Biol.* **39**, 221-244 (1988).
12. Bunya, M., Nishimura, M., Harashima, S. & Oshima, Y. *Molec. cell. Biol.* **11**, 3229-3238 (1991).



non-typical open-closed transitions. The authors estimate that the single-channel conductance is low, but that channel density in the peribacteroid membrane is high ( $300\text{--}2,400\ \mu\text{m}^{-2}$ ). Most importantly, they show that total ammonium currents carried by the  $\text{NH}_4^+$ -permeable channels can easily accommodate physiological rates of ammonium uptake into the plant. The identification of the  $\text{NH}_4^+$  uptake channel in the peribacteroid membrane will allow further characterization of its proposed physiological function and regulation mechanisms (for example, block by calcium<sup>2</sup>), and for comparisons with other cation transporters to be made.

The biophysical and cell biological characterization, as well as the cloning, of

initial cDNAs encoding key members of the NPK macronutrient transporters are now allowing accurate analyses of numerous physiological questions. For example, interactions of transporters with toxic compounds in soils, such as herbicide uptake by the low-affinity *CHL1* nitrate transporter<sup>3</sup>, or the interactions of toxic cations with cation-uptake transporters, are now open to investigation. We are no doubt in for surprises as we find out more about plant macronutrient uptake and the associated environmental stresses. □

*Julian I. Schroeder is in the Department of Biology and the Center for Molecular Genetics, University of California San Diego, La Jolla, California 92093-0116, USA.*

HIV

## An elusive soluble suppressor

*Anthony S. Fauci*

THE study by Baier *et al.* on page 563 of this issue<sup>1</sup> describes the identification of a soluble factor (lymphokine) secreted by  $\text{CD8}^+$ -bearing T lymphocytes that inhibits the replication of the human immunodeficiency virus (HIV). The potential implications of this study are substantial.

A variety of classical antibody and cell-based immune responses have been described that contribute to anti-HIV immunity, including neutralizing antibodies, antibody-dependent cellular cytotoxicity, natural killer cells, and HIV-specific cytotoxic T cells (reviewed in ref. 2). Several years ago, a phenomenon of non-cytolytic suppression of HIV replication was described<sup>3</sup>: the suppression was shown to be due to a soluble factor or factors derived from  $\text{CD8}^+$  T lymphocytes of HIV-infected individuals, and these findings were subsequently corroborated by several laboratories (reviewed in ref. 4). Similar phenomena and factors have been described in the simian immunodeficiency virus (SIV) model of HIV infection in the monkey<sup>5</sup> (also reviewed in ref. 4).

Despite continued attempts to isolate the factor(s), however, their precise identification has remained elusive. It is likely that the difficulties can be explained at least in part by the facts that there may be a family of suppressor factors, and that attempts to purify and assay just one of them have resulted in marked decreases and inconsistencies in suppressor activity compared to that of the crude supernatants. This may very well be the case with the human factor cloned by Baier and colleagues, which gives lower levels of suppression than those seen in studies on crude supernatants<sup>3-5</sup>.

Baier *et al.* identified a factor that had been previously described<sup>6</sup> in a different context from HIV. The factor has been

designated lymphocyte chemoattractant factor (LCF) or IL-16, and it suppresses HIV replication in  $\text{CD8}^+$  T-lymphocyte-depleted human peripheral blood mononuclear cells<sup>1</sup>. The current authors had previously identified an "immunodeficiency-virus-suppressing lymphokine" (ISL) that was secreted by  $\text{CD8}^+$  T cells of African green monkeys (AGM) and which potently suppressed the replication of SIV in monkey cells as well as the replication of HIV in human  $\text{CD4}^+$  T lymphocytes<sup>5</sup>. Baier *et al.* note that the sequences encoding human IL-16 and AGM IL-16 are very alike, in that the IL-16 genes of humans differ from those of African green monkeys in only 16 of their 390 coding nucleotides, resulting in only seven non-clustered amino-acid changes.

It is not surprising, despite the intensive and hitherto unsuccessful search for the identity of the soluble suppressor factor(s) of HIV replication, that at least one of the factors is a lymphokine with other well-known functions apart from suppression of HIV replication. In this regard, it is highly likely that other discrete factors with different non-HIV related functions will be described that also suppress the replication of HIV. This might lead to understandable confusion as to the true identity of the suppressor factor; in fact, it is likely that a combination of factors will be required for optimal suppression.

It is out of the question that soluble factor(s) have emerged in the host species specifically to suppress HIV replication. Rather, it is conceivable that the virus has evolved to use existing factors of its host to its own advantage. Because the suppression of virus replication by the newly cloned factor is non-cytolytic (that is, it does not involve host destruction of infected host cells), a consequence might

be the establishment of latency, with the resulting implication of prolonged viral infection without killing the host.

Although  $\text{CD8}^+$  T-lymphocyte-derived LCF (IL-16) is a natural  $\text{CD4}$  ligand<sup>7</sup>, the ways in which it inhibits HIV replication are unknown. It is likely that those mechanisms will be quickly elucidated as investigators come to grips with the cloned factor. It will be important to determine the effect of the factor on replication of primary isolates of HIV, as well as on *in vitro* endogenous replication of virus in cells from HIV-infected individuals.

Work on the  $\text{CD8}^+$  T-lymphocyte-suppressor phenomenon has been seriously impeded by the unavailability of purified factors, necessitating the use of crude or only partially purified supernatants. In essence, the new study<sup>1</sup> opens up fresh avenues of research into pathogenesis, treatment and vaccine development. Of particular importance is the possibility of a novel form of 'immune-based' therapy. Transplantation of tissue and reinfusion of *ex vivo* expanded autologous immune-competent cells into HIV-infected individuals have met with the difficulties common to any such approach, including tissue rejection, graft-versus-host reactions and unforeseen toxicities (reviewed in ref. 8). The availability of a soluble factor with demonstrable ability to suppress HIV replication will certainly expand the options for experimental therapy. It is particularly relevant that there is considerable homology between human IL-16 and AGM IL-16, adding importance to the SIV monkey model of HIV infection.

Of course, there is no guarantee that such therapy would be effective, and more than a decade of experience with HIV has taught us to be conservative in our projections of success until properly designed and controlled clinical trials have been performed. However, at the very least, the availability of this and other HIV-suppressive factors alone or in combination will tell us a great deal about HIV pathogenesis, and in this respect will provide the framework for the development of strategies for therapy and vaccine development. □

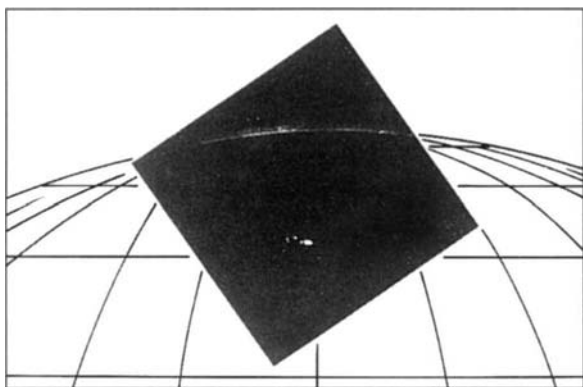
*Anthony S. Fauci is at the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.*

1. Baier, M., Werner, A., Bannert, N., Metzner, K. & Kurth, R. *Nature* **378**, 563 (1995).
2. Pantaleo, G., Graziosi, C. & Fauci, A. S. *New Engl. J. Med.* **328**, 327-335 (1993).
3. Walker, C. M. *et al. Science* **234**, 1563-1566 (1986).
4. Mackewicz, C. & Levy, J. A. *AIDS Res. Hum. Retrovir.* **8**, 1039-1050 (1992).
5. Ennen, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **91**, 7207-7211 (1994).
6. Cruikshank, W. W. *et al. Proc. natn. Acad. Sci. U.S.A.* **91**, 5109-5113 (1994).
7. LaBerge, S. *et al. J. Immun.* **155**, 2902-2910 (1995).
8. Fauci, A. S. *Science* **262**, 1011-1018 (1993).

# Probing times for Jupiter

Andrew P. Ingersoll

ON 7 December 1995, if all goes to plan, the Galileo spacecraft will swing into orbit around Jupiter, turning its back on the first-ever probe dropped into the planet's atmosphere. The probe, descending by parachute, should be able to take atmospheric data from above the tops of the clouds, where the pressure is several hundred millibar, down to a pressure of 25–30 bar, which is well below cloud base. Water is the key compound, both for the clues it contains about Solar System formation,



Lightning flashes (below) and auroral arc (above) on the dark side of Jupiter as seen by the Voyager 1 spacecraft in March of 1979.

and for its role in meteorology. On the Earth, moist convection is probably the single most important and least understood process in the climate system. Galileo will at last give us the chance to study moist convection on another planet.

Lightning was detected by the earlier Voyager spacecraft in images of the dark side of Jupiter<sup>1</sup> and in electromagnetic waves emanating from the planet. So the Galileo probe carries a lightning detector amidst its complement of instruments that measure temperature, pressure, clouds, chemical composition, solar and infrared radiation, and winds. All bear on the abundance of water and the process of moist convection.

Until Galileo starts to supply the answers, modelling by extension of what we know to occur on the Earth is our best hope of understanding the climate of other planets. Two recent papers, one on page 592 of this issue, illustrate this: Yair *et al.*<sup>2</sup> and Gibbard *et al.*<sup>3</sup> apply knowledge of the Earth's atmosphere to calculate lightning in a jovian thundercloud. Heat from below drives rapid vertical motions, which carry cloud particles upwards. Charge transfer occurs when ice particles collide with partly frozen hail. Large-scale electrification follows as the small particles are swept upwards, leaving the large particles behind. A key point is that the modelled clouds produce lightning only

when the water abundance exceeds a certain critical value which is close to 'solar' composition — what you would get by cooling a piece of the Sun down to jovian temperatures. Such an atmosphere would be 99 per cent hydrogen and helium. Water would be the most abundant compound, as oxygen is the next most abundant element on the Sun. Will the Galileo probe bear this out, and if so, what are the implications?

Water controls the stability of a climate system — how resistant the atmosphere is to convective overturning. If the atmosphere has abundant water and high stability, the weather might be confined to a layer less than 100 km thick within the jovian clouds. If the atmosphere has no water and low stability, the weather might extend 10,000 km or more into Jupiter's fluid interior. Such differences may help clarify an enduring jovian mystery — how storms like the Great Red Spot can last for hundreds of years whereas storms on the

Earth last for days or weeks. Although the winds are stronger than on the Earth, the weather on Jupiter is more predictable. Galileo should help us understand why.

The amount of water on Jupiter has remained a mystery because its vapour pressure is low at the tops of the jovian clouds. Last year, waves from the impacts of fragments of comet Shoemaker–Levy 9 provided indirect evidence that the atmosphere is richer in water than solar composition<sup>4</sup>, but the Galileo probe provides the definitive test. Apart from casting light on atmospheric motion, the water abundance will help in assessing the role of comets in distributing volatiles in the early Solar System. A high oxygen abundance relative to carbon and nitrogen suggests a large role for comets during the formation, not just of Jupiter, but of the Earth as well. Planetary scientists will be eagerly awaiting the first data from the probe. □

Andrew P. Ingersoll is in the Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA.

1. Borucki, W. J. & Williams, M. A. *J. geophys. Res.* **91**, 9893–9903 (1986).
2. Yair, Y., Levin, Z. & Tzivion, S. *Icarus* **115**, 421–434 (1995).
3. Gibbard, S., Levy, E. H. & Lunine, J. I. *Nature* **378**, 592–595 (1995).
4. Ingersoll, A. P. & Kanamori, H. *Nature* **374**, 706–708 (1995).

## Reformed rubbish

BIOLOGY has no waste products. All exhaled or excreted matter, and every dead thing, is raw material for some other organism. It can be hydrolysed back to basic biomolecules and used again. Industrial society has no such general recycling reaction. Burning is a poor substitute: it breaks down and detoxifies most industrial and consumer rubbish, but yields only smoke and ash in return. Daedalus now proposes a true universal recycler — steam reforming.

Modern industrial steam reforming is conducted, often catalytically, on hydrocarbons. Hot, high-pressure steam converts them to hydrogen, methane and carbon oxides, useful feedstock gases. The process should work on almost any organic material. Daedalus now reckons that nature got there first. He recalls the mysterious geological process of petrification, which slowly replaces dead plants and animals by silica copies of wonderful fidelity. Much of our knowledge of ancient life forms comes from their petrified remains.

The process is badly understood. Normal groundwater contains far too little silica to replace the organic bulk of (say) a tree trunk. One theory is that petrification occurs with hot water, which can carry more silica; or even steam, which surprisingly can dissolve still more. (Silica dissolved in steam is a major headache for turbine engineers.)

So Daedalus reckons that petrification is simply geological steam reforming. Hot, silica-saturated steam from local geothermal action diffuses into the dead specimen and reforms it. The resulting gases do not dissolve silica, which is precipitated. The more organic material at any point of the specimen, the more steam is used up in the reaction, and the more silica is precipitated. The organic structure is thus perfectly replaced by silica. The process could be very fast — many vapour-phase cracking reactions use silica-based catalysts.

This elegant theory suggests a new form of rubbish disposal. Daedalus hopes to replace our unpopular incinerators with giant silicated-steam reformers for all kinds of organic and plastic junk. Old newspapers, discarded clothes and furniture, packaging and food waste, obsolete computers, all will be catalytically steam-reformed away to useful gases, leaving perfect silica copies in their place. These should really be crushed and fed back into the steam for recycling; but Daedalus likes the idea of burying them as landfill, to astonish future archaeologists. He also plans a steam reforming crematorium, which simultaneously produces a silica statue of the deceased.

David Jones

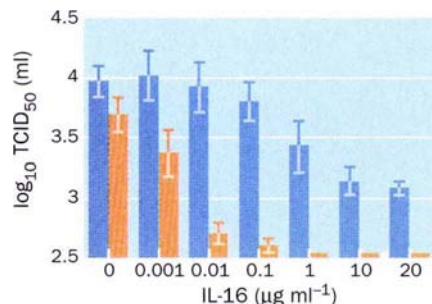
# HIV suppression by interleukin-16

**SIR** — In addition to their function as cytotoxic T cells, lymphocytes carrying the CD8 receptor (CD8<sup>+</sup>) secrete one or more soluble factors which inhibit replication of HIV and SIV (human and simian immunodeficiency viruses, respectively) in primary CD4<sup>+</sup> cells<sup>1,2</sup>. This antiviral activity is positively correlated with the health of HIV-infected patients, and is high during the asymptomatic stages of infection. Old World monkeys such as the African green monkey *Cercopithecus aethiops*, which are naturally infected with SIV in the wild, possess high immunodeficiency virus-suppressing lymphokine (ISL) activity<sup>2</sup>, a low viral load<sup>3</sup>, and never develop simian AIDS<sup>4</sup>.

Here we show that the previously identified lymphocyte chemoattractant, interleukin-16 (IL-16), which is secreted from activated CD8<sup>+</sup> cells and binds to T cells through the CD4 receptor<sup>5,6</sup>, suppresses the replication of HIV and SIV. The IL-16 from African green monkeys is highly homologous to its human counterpart. It may therefore contribute to the low viral load in healthy HIV-infected patients and naturally infected African green monkeys.

The IL-16 molecules isolated from African green monkeys and humans (Fig. 1) are proteins of 130 amino acids and relative molecular mass around 13,500. The genes encoding human (HU) and African green monkey (AGM) IL-16 differ in 16 of their 390 coding nucleotides, resulting in 7 non-clustered amino-acid changes.

On analysis by gel filtration, recombinant soluble IL-16<sub>HU</sub> and IL-16<sub>AGM</sub> elute predominantly as homodimers (apparent  $M_r$  28,000), with minor peaks corresponding to monomeric and tetrameric forms. The lymphocyte chemoattractant activity of IL-16 was previously linked to its tetramer formation<sup>5</sup>, so it is possible that a



**FIG. 2** Inhibition of HIV-1 replication by human (blue columns) and simian (pink columns) recombinant IL-16 (rIL-16) on CD8<sup>+</sup>-depleted PBMCs. CD8<sup>+</sup>-depleted PBMCs ( $1.5 \times 10^6$ ) stimulated with phytohemagglutinin (CD8<sup>+</sup> cells <5%, by fluorescence-activated cell sorting) were plated in RPMI 1640 medium supplemented with 20% FCS, 2 mM glutamine and 180 U ml<sup>-1</sup> IL-2 into 96-well plates and infected with 50 TCID<sub>50</sub> (the 50% tissue-culture infectious dose) of PBMC-grown HIV-1<sub>SF2</sub> for 1 h. Unbound virus was removed, and cells were replated in culture medium containing IL-16. At 8, 10 and 12 days post-infection, virus content of cell-culture supernatants was determined by titration on the highly HIV-1<sub>SF2</sub>-susceptible cell line MT-4 (NIH AIDS Research and Reference Reagent Program). TCID<sub>50</sub> was calculated according to Spearman/Kärber. The result of titrations of quadruplicates 10 days post-infection are shown.

proportion of the IL-16<sub>HU</sub> and IL-16<sub>AGM</sub> preparations contain incorrectly folded protein. Dimerization of the human IL-16 is partially due to disulphide bridging between the cysteine residues at position 31, which is replaced by tyrosine in IL-16<sub>AGM</sub>. Neither thrombin treatment to remove the histidine tag of the fusion proteins nor changes in induction by isopropyl-β-D-thiogalactoside or bacterial growth temperature abolish the high degree of dimerization (data not shown). Higher concentrations of IL-16<sub>HU</sub> than of IL-16<sub>AGM</sub>

are needed to achieve inhibition of HIV replication (Fig. 2). Preliminary data suggest that the apparent differences in activity between IL-16<sub>HU</sub> and IL-16<sub>AGM</sub> are due to different folding characteristics of the bacterially derived proteins.

Although IL-16 has been shown to induce signal transduction<sup>5,7</sup>, the precise way in which it inhibits HIV replication is unknown. Within the concentration range tested, IL-16 does not impair cell viability as judged by trypan blue exclusion and cell number determination. It is known to bind to the CD4 receptor, and so may resemble certain anti-CD4 antibodies that inhibit transcription driven

by the long terminal repeats of HIV without affecting initial binding by the virus<sup>8</sup>. Further investigation is needed to determine what fraction of the ISL activity is due to IL-16, but it has not escaped our attention that IL-16 may possess an antiviral therapeutic effect.

**Michael Baier, Albrecht Werner  
Norbert Bannert, Karin Metzner  
Reinhard Kurth**

Paul-Ehrlich-Institut,  
Paul-Ehrlich-Strasse 51–59,  
D-63225 Langen, Germany

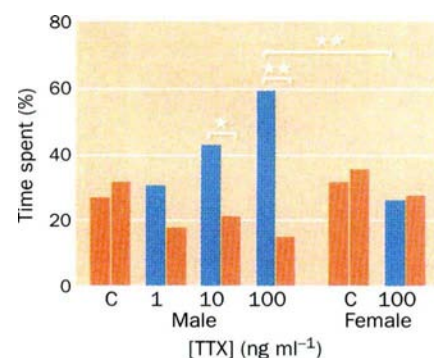
- Levy, J. A. *AIDS* **7**, 1401–1410 (1993).
- Ennen, J. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 7207–7211 (1994).
- Hartung, S. et al. *J. Virol.* **66**, 2143–2149 (1992).
- Norley, S. G. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 9067–9071 (1990).
- Cruikshank, W. W. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 5109–5113 (1994).
- Laberge, S. et al. *J. Immun.* **155**, 2902–2910 (1995).
- Cruikshank, W. W. et al. *J. Immun.* **138**, 3817–3823 (1987).
- Corbeau, P. et al. *J. Immun.* **150**, 290–301 (1993).

M.B. and A.W. contributed equally to this study.

## Tetrodotoxin as a pheromone

**SIR** — Tetrodotoxin (TTX) is one of the most potent marine toxins<sup>1</sup>, and so one possible physiological role is as a defence agent<sup>2,3</sup>. Here I report another function of TTX, as a male-attracting pheromone at the time of spawning.

Measurement of attracting activity of the puffer fish *Fugu niphobles* against TTX was conducted by using a Y-maze



**FIG. 1** Preference activities of *F. niphobles*, against TTX. Results represent medians of eight males and females in each concentration. \* $P < 0.05$ ; \*\* $P < 0.01$  when compared with control reactions. The preference reactions of spermiating males and sexually mature females were tested by the Y-maze flow-through system<sup>4</sup> operated after a fish was placed in the starting point at the centre of Y maze. Artificial sea water was introduced into both arm inlets of the maze (dimensions  $45 \times 8 \times 8$  cm) with a flow of 2,000 ml min<sup>-1</sup>. The TTX solution and control distilled water were introduced into the arm channel in each inlet with a constant flow of 1.0 ml min<sup>-1</sup>. The percentage of the spent period (5 min) for the test fish in each channel was analysed statistically with Friedman's two-way analysis of variance, followed by Wilcoxon's signed-rank test or the Mann-Whitney *U*-test. Blue columns, TTX; pink columns, distilled water.



**FIG. 1** a, Nucleotide sequence of IL-16<sub>AGM</sub>. Differences from the human IL-16 sequence are underlined. Codon 26, which is deleted in some of the human and African green monkey IL-16 alleles, is marked by a double line. b, Alignment of the deduced amino acid sequences of IL-16<sub>HU</sub> and IL-16<sub>AGM</sub>. Peripheral blood mononuclear cells (PBMCs;  $5 \times 10^6$ ) were cultivated with 10 µg ml<sup>-1</sup> concanavalin A and IL-2 for 48 h. Total RNA was prepared and a cDNA synthesis was performed in a standard polymerase chain reaction. Purified products were cloned for nucleotide sequence confirmation and protein expression into the vector pET15b (Novagen). Expressed IL-16 was obtained by Ni<sup>2+</sup>-NTA agarose (Qiagen) chromatography to >95% purity.



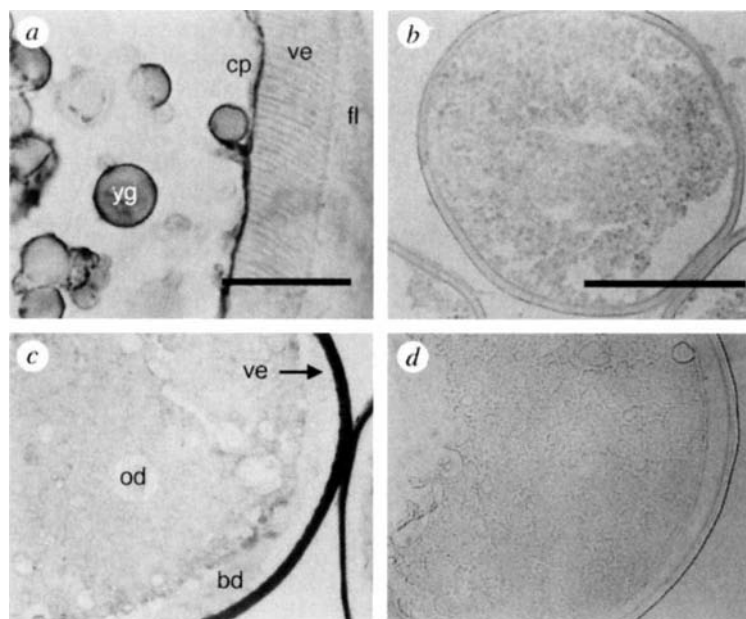


FIG. 2 Immunocytochemical localization of TTX in the vitelline envelope of pre- and post-ovulated oocytes. *a*, Pre-ovulated oocyte in the ovigerous lamella; *b*, negative control of *a*; *c*, ovulated oocyte; *d*, negative control of *c* (yg, yolk globule; cp, cytoplasm; ve, vitelline envelope; fl, follicle layer; od, oil droplet). Scale bars: *a*, 40  $\mu$ m; *b*, 200  $\mu$ m. The cryostat sections (5–7  $\mu$ m) of neutral formaldehyde-fixed ovulated oocytes and intraovarian pre-ovulated oocytes were stained by a monoclonal antibody against TTX<sup>5</sup> (2.5  $\mu$ g ml<sup>-1</sup>). Indirect immunocytochemistry was performed using the reagents and suggested protocols of Vectastain ABC kits (Vector). As additional negative controls, a monoclonal antibody pre-absorbed with TTX was applied to the section.

flow-through system<sup>4</sup>. TTX had an attracting pheromonal activity to spermating male *F. niphobles*, but not to the sexually active females (Fig. 1). Males could recognize 10 ng of TTX in 2,000 ml, indicating that 15 pM of TTX is enough to attract fishes and that the minimum effective concentration lies in the range 15–1.5 pM.

To demonstrate that TTX is released from the ovulated oocytes, five females of *F. niphobles* were caught during the spawning season (June 1995) and both pre- and post-ovulated oocytes collected in each ovary. The TTX concentration per oocyte was measured by an indirect competitive enzyme immunoassay using a monoclonal antibody against TTX<sup>5</sup>. TTX levels in ovulated oocytes was markedly less than in pre-ovulated oocytes ( $P < 0.05$ ) and about half of the toxin (23.9 of 41.9 ng per oocyte on average) was lost from the oocytes after ovulation. Immunocytochemical staining of the intra-ovarian pre-ovulated oocytes showed that TTX was present in the limiting membrane of yolk

globules and the cytoplasm, while no evidence of TTX was obtained in the vitelline envelope (Fig. 2*a*). In contrast, the vitelline envelope of the ovulated oocytes was intensely stained (Fig. 2*c*). These observations strongly suggest that TTX is released from the oocytes through the vitelline envelope during ovulation, probably as a result of biochemical changes in the oocytes at the time of germinal vesicle breakdown and hydration.

Various pheromonal attractants, such as prostaglandins<sup>6</sup> and steroids<sup>7</sup>, have been chemically identified in teleosts<sup>8</sup>. They accumulate and are then discharged in the breeding season. *F. niphobles* form schools and rush to beaches for spawning, raising the possibility that TTX facilitates formation of these breeding aggregations. Also, it is well known that TTX levels in puffer fish ovaries is elevated with ovarian maturation in the breeding season<sup>1</sup>. Ovulated *F. niphobles* oocytes are transparent and their volume increases fourfold over pre-ovulated oocytes, suggesting that significant biochemical and physiological changes may occur during ovulation. In fact, the decrease and transition of TTX in oocytes observed in the course of ovulation, shows that after TTX is released from ovulated oocytes into the ovarian cavity, it is discharged from the cloaca before spawning.

This study shows that TTX has a clear male-attracting function and that TTX is released from ovulated oocytes through

the vitelline coat. This further suggests that similar substances with low relative molecular mass may be released at the time of ovulation in other teleosts and vertebrates and may play a physiological role in spawning or fertilization.

**Kendo Matsumura**

Yamaguchi Prefectural Research  
Institute of Health,  
2-5-67 Aoi Yamaguchi 753, Japan

## Ovalocytosis and cerebral malaria

**SIR**—The geographical coincidence of South-East Asian ovalocytosis and malaria suggests that this condition might represent a red blood cell variant conveying protection against malaria<sup>1</sup>. The demonstration of reduced parasitaemia in heterozygous individuals<sup>1–4</sup> supports this hypothesis, but does not prove an advantage conferred by South-East Asian ovalocytosis. To test the hypothesis that children with ovalocytosis are protected against death from malaria, we have compared the prevalence of the 27-base-pair deletion in the band 3 (AE1) gene that causes the condition, in populations with different clinical status of malaria in the Madang area, Papua New Guinea.

There is a clear pattern of decreasing prevalence of band 3 deletion with increasing disease severity. None of the cerebral malaria cases but 15% of the healthy individuals had a deletion of the band 3 gene; uncomplicated malaria had a prevalence rate of 9%. Among healthy individuals, those without *Plasmodium falciparum* parasitaemia had a higher prevalence of band 3 deletion than those with parasitaemia. The negative association between band 3 gene deletion and cerebral malaria was especially strong after allowance was made for factors that were unequally distributed between the groups such as age and geographical area (see table).

The mechanism of malaria resistance

1. Baer, A., Lie-Injo, L. E., Welch, Q. B. & Lewis, A. N. *Am. J. hum. Genet.* **28**, 179–188 (1976).
2. Serjeantson, S., Bryson, K., Amato, D. & Babona, D. *Hum. Genet.* **37**, 161–167 (1977).
3. Cattani, J. A., Gibson, F. D., Alpers, M. P. & Crane, G. G. *Trans. R. Soc. trop. Med. Hyg.* **81**, 705–709 (1987).
4. Foo, L. C., Rekhraj, V., Chiang, G. L. & Mak, J. W. *Am. J. trop. Med. Hyg.* **47**, 271–275 (1992).
5. Kidson, C., Lamont, G., Saul, A. & Nurse, G. T. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5829–5832 (1981).
6. Hadley, T., Saul, A., Lamont, G. & Hudson, D. E. *J. clin. Invest.* **71**, 780–782 (1983).
7. Crandall, I. & Sherman, I. W. *Parasitology* **108**, 257–267 (1994).
8. Carlson, J., Nash, G. B., Gabutti, V., Al-Yaman, F. & Wahlgren, M. *Blood* **84**, 3909–3914 (1994).
9. Dobzhansky, T. in *The Structure of Human Populations* (eds Harrison, G. & Boyce, A. J.) 212–233 (Clarendon, Oxford, 1972).
10. World Health Organization *Trans. R. Soc. trop. Med. Hyg.* **84** (Suppl. 2), 1–65 (1990).
11. Al-Yaman, F. et al. *Trans. R. Soc. trop. Med. Hyg.* **89**, 55–58 (1995).
12. Jarolim, P. et al. *Proc. natn. Acad. U.S.A.* **88**, 1022–1026 (1991).

1. Kao, C. Y. *Pharmac. Rev.* **18**, 997–1049 (1966).
2. Fuhrman, F. K. *Ann. N.Y. Acad. Sci.* **479**, 1–14 (1986).
3. Kodama, M., Ogata, T. & Sato, S. *Mar. Biol.* **87**, 199–202 (1985).
4. Ishida, Y. & Kobayashi, H. *Fish. Sci.* **61**, 441–446 (1995).
5. Matsumura, K. *Eur. J. Pharmac.* **239**, 41–45 (1995).
6. Kobayashi, M. & Stacey, N. *Horm. Behav.* **27**, 38–55 (1993).
7. Nagahama, Y. in *Hormones and Reproduction in Fishes, Amphibians and Reptiles* (eds Norris, D. O. & Jones, R. E.) 171–202 (Plenum, New York, 1987).
8. Dulka, J. G., Stacey, N. E., Sorensen, P. W. & Van Der Kraak, G. *J. Nature* **325**, 251–253 (1987).

## OCCURRENCE OF BAND 3 DELETION AMONG CHILDREN IN THE MADANG AREA OF PAPUA NEW GUINEA

|                                                                                | Cerebral malaria (CM)<br><i>n</i> =35 |         | Uncomplicated malaria (UM)<br><i>n</i> =57 |         | Healthy individuals (HI)<br><i>n</i> =103 |
|--------------------------------------------------------------------------------|---------------------------------------|---------|--------------------------------------------|---------|-------------------------------------------|
| <b>Demographics and parasitology</b>                                           |                                       |         |                                            |         |                                           |
| Mean age in years (s.d.)                                                       | 4.2 (2.7)                             | <0.001* | 3.5 (1.9)                                  | <0.001† | 5.8 (2.1)                                 |
| Range                                                                          | 1–10                                  |         | 0.8–11                                     |         | 2–11                                      |
| Area: North (%)                                                                | 19 (54) in 12 villages                | 0.20*   | 35 (61) in 22 villages                     | 0.02†   | 43 (42) in 11 villages                    |
| South (%)                                                                      | 16 (46) in 13 villages                |         | 22 (39) in 15 villages                     |         | 60 (58) in 10 villages                    |
| Mean log <sub>10</sub> ( <i>P. falciparum</i> +1)<br>(parasites per µl (s.d.)) | 4.05 (1.25)                           | <0.001* | 4.09 (0.68)                                | <0.001† | 2.04 (1.71)                               |
| <b>Gene frequency and morbidity association</b>                                |                                       |         |                                            |         |                                           |
| Band 3 deletion<br>prevalence rate and CI                                      | 0% 0–10%                              |         | 8.8% 3–19%                                 |         | 14.6% 8–23%                               |
| Odds ratio, CI and <i>P</i> value                                              |                                       |         |                                            |         |                                           |
| Unadjusted                                                                     | 0 0–0.75                              | 0.01*   | 0.56 0.15–1.76                             | 0.33    | 1                                         |
| Adjusted for age and<br>geographical area                                      | 0 0–0.21                              | <0.001* | 0.26 0.06–0.94                             | 0.04    | 1                                         |

Only children living in villages are included. Those living in towns or in suburban areas have different ethnic origins, with known differences in prevalence of the deletion. CM, patients recruited at the Madang General Hospital, defined according to published criteria<sup>10</sup>, with a slight modification of the Blantyre coma score<sup>11</sup>. UM, children who attended health facilities in the same areas with a recent history of fever, no other major symptoms or signs, and with a confirmed *P. falciparum* asexual blood stage parasitaemia. HI, children who had not complained of symptoms during the previous week. They were included irrespective of their parasitological status. Deletion in the erythrocyte band 3 gene was determined using polymerase chain reaction (see, for example, ref. 12). The strength of association between the band 3 deletion and disease severity was assessed by comparing the prevalence rate of the red-cell variant gene in CM cases and in HI using Fisher's exact test (two-tailed), and by the exact 95% confidence interval (CI) for the odds ratio. UM and HI were also compared. Adjustment for age and area was made by logistic regression. A quadratic effect of age was fitted because of a significant departure from linearity. Likelihood-based confidence intervals were calculated. \* *P* value for CM against HI; † *P* value for UM against HI.

in South-East Asian ovalocytosis is speculative. It is known that ovalocytes are relatively resistant to invasion<sup>5,6</sup> and that they provide some protection against parasitaemia<sup>3</sup>. It seems likely from this study that ovalocytosis might protect against uncomplicated morbidity from malaria to about the same extent as it protects against parasitaemia. Cerebral malaria is a potentially fatal disease, so the frequency of the band 3 deletion in these patients can be taken as an indicator of its selective advantage. We could find no heterozygous individuals among the cerebral malaria cases; ovalocytosis would thus seem to protect from death by malaria.

The view that ovalocytosis is maintained in the population by an advantageous effect operating solely against parasite invasion is questionable. Experimental observations suggest that band 3 deletion plays a specific role in the pathogenesis of cerebral malaria through decreased cytoadherence of infected red cells to cerebral microvessels. This process follows a modification of the band 3 protein in infected red cells and cytoadherence can be blocked in mice with polyclonal antiserum against a polypeptide corresponding to the human band 3 residues 546–553 (ref. 7). Unlike the case of thalassaemic and HbS-containing red cells, previous studies, including samples from the Madang area, have not supported a modifying role of ovalocytic cells on the rosetting ability of red cells<sup>8</sup>. It seems unlikely, therefore, that the protection against cerebral malaria conferred by ovalocytosis in Papua New Guinea operates through the rosetting mechanism.

In screens of around 2,000 subjects (including heterozygous matings) (C.S.M. *et al.*, unpublished observations), we have found no individuals homozygous for band 3 deletion. It is thus likely that homozygosity is lethal *in utero*. Ovalocytosis in Papua New Guinea may not be a balanced polymorphism. However, if we assume an equilibrium and that all homozygotes for the band 3 deletion die *in utero*, a prevalence of 15% heterozygosity in the general population, as we have

in this study, would imply that 9% of all homozygotes without the gene variant die of malaria before reproduction<sup>9</sup>.

**Blaise Genton, Fadwa Al-Yaman  
Charles S. Mgone, Neal Alexander  
Michael M. Panlu, Michael P. Alpers**  
Papua New Guinea Institute of  
Medical Research,  
PO Box 378, Madang, Papua New Guinea  
**David Mokela**  
Madang General Hospital,  
Madang, Papua New Guinea

## Extrapolation or attention shift?

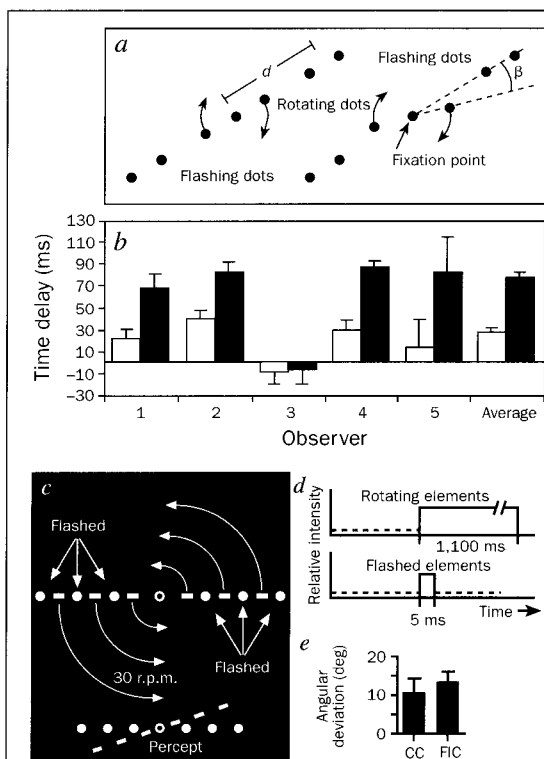
**SIR** — Interceptive actions, such as hitting a moving target or catching a ball, need to be adequately timed in order to be successful. Because there is a significant time delay in the transmission of information along the visual pathways, there could be a critical difference between the perceptual and actual position of a moving object. Nijhawan<sup>1</sup> proposed that a moving object is seen closer to its actual physical location, extrapolated from features of the motion, partially compensating for the delay in visual processing.

Most observers report a perceptual misalignment between two sets of aligned dots when one set is in motion (*a* in the figure). Two outer pairs of dots, diametrically opposed to each other, are momentarily flashed in perfect alignment with an inner pair of continuously rotating dots. The moving dots, which are generally seen ahead of the flashing dots, would have their position, according to Nijhawan<sup>1</sup>, perceptually extrapolated by about 80 ms.

We interpret this effect as resulting from

a longer time delay involved in the visual processing of the flashing dots, rather than an extrapolation of the moving dots. In a series of experiments to investigate attentional mechanisms, we found a clear dependence of this effect on the location of the flashing dots. The magnitude of the perceived misalignment increased as the flashing dots were moved away from the fixation point (*b* in the figure).

In further, but still preliminary, experiments with two other observers, using a similar set-up, we also examined the case in which the positions of the flashing and moving dots were switched, as well as the original experiment. In all situations, the moving dots were perceived as located ahead of the flashing dots. However, in the switched version of the experiment, the mean angle of perceived misalignment showed a much smaller, or no, dependence on the location of the moving (outer) dots. Furthermore, its magnitude was comparable to those obtained when the flashing (outer) dots were closer to the



Experimental design and results of: *a, b*, Baldo and Klein, and *c-e*, Khurana and Nijhawan. *a, Left*, the stimulus, a pair of dots  $1.3^\circ$  apart in the visual field, rotating at 25 r.p.m. about a central dot (fixation point). Two pairs of dots were flashed in perfect alignment with the central and rotating dots. Distances, *d*, from the central dot to the midpoint between each pair of flashing dots, used in our tests, were  $1.45^\circ$  and  $4.74^\circ$ . *Right*, The situation as reported by most observers: the rotating dots are seen ahead of the flashing dots. The perceived misalignment,  $\beta$ , was assessed by letting the observer adjust the flashing dots until they appeared in alignment. *b*, The misalignment as a time delay and the mean time delay from five naive observers measured at two different distances (data averaged over all observers and weighted by inverse variance). Two other observers, under a similar experimental set-up, offered similar results concerning the original experiment (average time delays:  $33 \pm 9$  ms and  $83 \pm 7$  ms for closer and farther flashing (outer) dots, respectively). A switched version of the experiment was also used where the outer dots moved and the inner dots flashed. The perceptual effect was qualitatively the same, but the strong dependence on the location of the outer (moving) dots was not observed (average misalignment:  $24 \pm 14$  ms and  $19 \pm 9$  ms for the closer and farther locations, respectively). These values are comparable to those found in the original experiment, when the flashing (outer) dots were closer to the fixation point.  $\square$ ,  $1.45^\circ$ ;  $\blacksquare$ ,  $4.74^\circ$ . *c*, The observer fixated the central dot while attentively tracking a line (length =  $6.9^\circ$ ) composed of 6 rectangles rotating at 30 r.p.m. A horizontal line (length =  $7.8^\circ$ ), composed of 6 circles interleaving the rotating line, was flashed for 5 ms. Observers reported the perceived relative positions of the two lines by varying the angle between two adjustable lines. A strong flash-lag effect was reported (percept). *d*, The intensity profiles of the rotating and flashed elements as a function of time in the FIC condition. *e*, Of 10 observers, 5 viewed the FIC condition first and the complete cycle (CC) second; the order was reversed for the remaining five. A paired *t*-test showed no significant difference between the angle means for the CC and FIC conditions; mean FIC - mean CC = 2.8,  $t(9) = 1.70$ ,  $P > 0.10$ . The average ( $n = 10$ ) angular deviations for the CC and FIC conditions yield time delays of 58.61 and 74.17 ms, respectively.

fixation point in the original experiment (see figure legend).

These findings support the idea that the perceptual effect is mainly involved in the location of the flashing dots in the visual field. We hypothesize that some amount of time, dependent on eccentricity, is required to bring the flashing dots to a sufficiently high level of sensory awareness for a 'snapshot' of the moving dots to be taken. Such a time delay would be related to the abrupt onset of the flashing dots and might involve attentional mechanisms, either in capturing attention<sup>2</sup> or in shifting the focus of attention from one place to another across the visual field<sup>3,4</sup>.

Purely sensorial mechanisms, operating preattentively and depending on eccentricity, cannot yet be discarded. More experiments are needed to distinguish between these possibilities, but we believe that an attentional hypothesis should be examined further.

**Marcus Vinicius C. Baldo**

Departamento de Fisiologia e Biofísica,  
Instituto de Ciências Biomédicas I,  
Universidade de São Paulo,  
São Paulo, São Paulo 05508, Brazil

**Stanley A. Klein**

School of Optometry,  
University of California at Berkeley,  
Berkeley, California 94720, USA

esis that the lag of predictably moving objects is 'corrected' by the visual system through extrapolation<sup>1</sup>, or through some form of neural facilitation applied along the inferred trajectory of moving objects. Neural responses for inferred motion, where an explicit motion signal from the retina is absent, have been observed<sup>5</sup>. Flashed objects, on the other hand, are unpredictable and subject to expected neural delays, which cause their apparent lag.

We addressed the 'attention shift' hypothesis by spatially interleaving the moving and flashed elements. In this condition, observers attentively tracked<sup>6</sup> a rotating line composed of six rectangles. A horizontal line composed of six circles was flashed for 5 ms (*c* in the figure). As the flashed elements occupy the spaces between the attended rotating elements, attention shifts should be negligible and the flash-lag effect should not be observed. However, this display produced a strong effect (*e* in the figure).

Delays due to 'attention capture' were tested by the abrupt onset of both the moving and flashed elements. The display was modified such that the flashed and rotating elements came on simultaneously for 5 and 1,100 ms, respectively. In this 'flash-initiated' cycle (FIC), the rotating and flashed elements have an equally abrupt onset (*d* in the figure), and thus both capture attention equivalently. If delays due to attention capture cause the effect, then none should be observed in this condition. The effect we found, however, did not significantly differ in strength from that observed in the

'complete' cycle (*e* in the figure). When flashed and moving objects are equated in terms of the shift-time or capture-time of attention, observers continue to report the flash-lag effect.

We explain the FIC results on the basis of parallel processing in the visual system<sup>7,8</sup>. The neural processing of both the rotating and the flashed lines begins simultaneously, but the observer does not perceive either stimulus for approximately 100 ms (ref. 9). During this time, the retinal image of the line rotating at 30 r.p.m. has moved through  $18^\circ$ , triggering a motion signal in the magnocellular stream. We suggest that lag-correction, which probably occurs in the fast magnocellular stream, is implemented within that period. The correction process is complete within the time window required for neural signals to yield visual awareness, and before the onset of attentional processing<sup>10</sup>.

**Beena Khurana, Romi Nijhawan**

Department of Psychology,  
Uris Hall, Cornell University,  
Ithaca, New York 14853, USA

1. Nijhawan, R. *Nature* **370**, 256-257 (1994).
2. Hillstrom, A. & Yantis, S. *Percept. Psychophys.* **55**, 399-411 (1994).
3. Tsai, Y. J. *exp. Psychol.* **9**, 523-530 (1983).
4. Weichselgartner, E. & Sperling, G. *Science* **238**, 778-780 (1987).
5. Assad, J. A. & Maunsell, J. H. R. *Nature* **373**, 518-521 (1995).
6. Cavanagh, P. *Science* **257**, 1563-1565 (1992).
7. Livingstone, M. S. & Hubel, D. H. *J. Neurosci.* **7**, 3416-3468 (1987).
8. Phillips, C. G., Zeki, S. & Barlow, H. B. *Brain* **107**, 328-360 (1984).
9. De Valois, R. L. & De Valois, K. K. *Vision Res.* **31**, 1619-1626 (1991).
10. Treisman, A. in *Attention: Selection, Awareness and Control* (eds Baddeley, A. & Weiskrantz, L.) 5-35 (Clarendon, Oxford, 1993).



# Chemistry imagined

Peter Atkins

**Force of Nature: The Life of Linus Pauling.** By Thomas Hager. *Simon and Schuster*: 1995. Pp. 721. \$35.

**Linus Pauling: A Life in Science and Politics.** By Ted Goertzel and Ben Goertzel. *BasicBooks*: 1995. Pp. 300. \$27.50.

**Linus Pauling in His Own Words: Selections from His Writings, Speeches, and Interviews.** Edited by B. Marinacci. *Simon and Schuster*: 1995. Pp. 320. \$27.50 (hbk); £15 (pbk).

LINUS Carl Pauling was born as the century opened, and died almost at its end. He lived through politically, socially and scientifically turbulent times, and contributed to them all. It is fitting that his life has been analysed for us, and helpful that two volumes have appeared independently, for his contributions have been controversial, and several eyes are needed to see them in the round.

Pauling was born in Portland, Oregon, in 1901, the grandson (on his father's side) of German immigrants, and his early life was overshadowed by the premature death of his father, a pharmacist, and the inability of his mother to show the warmth that we are told children should expect. Through the relative trivia of his childhood (but then whose childhood is other than trivial?) we see glimpses of emerging selfishness and self-centredness as well as a deep interest in the workings of the world. These strands were to form the mainsprings of his scientific life, just as compassion, unconventionality and hubris later joined forces to motivate his political life.

He certainly had to struggle in his early days, and perhaps acquired then the determination — or obstinacy — that was to be a hallmark of his life. His intellect started to be brought under control at Washington High School in Portland, where he first encountered systematic thinking. At about the same time, his life-long passion for chemistry was kindled when his friend demonstrated to him a few simple experiments, and from then on Pauling realized that things were not immutable. He later identified this moment as the start of his long career in chemistry.

The only college he could afford to attend was Oregon Agricultural College (now Oregon State University) at Corvallis, about 90 miles from his home. There was a tension between him and his mother, who wished him to get a real job and help contribute to the family's tiny

income. But his wishes prevailed, and Pauling put in long hours of study and paid employment to get himself through college, from which he graduated in

## IMAGE UNAMAGABLE FOR A CORN LABET FOR A CORN LABET REASONS

**In the air — one of Pauling's more controversial theories was that high concentrations of vitamin C can cure cancer.**

chemical engineering in 1922. Here again we see the seeds of his later life, for devotion to hard work would always be his style, right to the end. It was at this time that he became aware not merely of the mutability of matter but also that there would be an explanation in terms of, presumably, atoms and electrons. He later said that reading one of Irving Langmuir's papers had provided him with the crucial spark. He had caught sight of the underworld of explanation.

He fell in love, too. For now he was to meet his lifelong wife, Ava Helen Miller, a companion who in due course would have a crucial influence on his political views. She was entwined in his development, yet seemingly wholly subservient to

it. Pauling emerges as the master and, despite his unconventionality in so many fields, Ava Helen remained the conventional wife, attending to wifely duties as Pauling's star ascended, but willing to discard her mothering for long periods if it got in the way of their joint inclination to travel. Just occasionally we see a glimpse of her frustration; who knows the depths it actually reached; both biographies allude to the extent but do not explore it.

Next, to the California Institute of Technology, the institution that would be Pauling's intellectual base for more than 40 years and which would increasingly find his presence uncomfortable before old age began to mellow memories, death removed his enemies and room was made for reconciliation. Pauling arrived in Pasadena at the start of Caltech's blossoming into greatness, and undoubtedly repaid the heavy debt he owed it by the time he came to leave. "King Arthur" Noyes was responsible for building up chemistry, and the institute had just induced America's best-known physicist, R. A. Millikan, to migrate there from Chicago to help establish its reputation in physics. Here was exactly the intellectually stimulating environment, encouraging the crossing of disciplinary borders, that would provide Pauling with the perfect compost for his mould of science.

Important formative influences took Pauling abroad between 1925 and 1927, particularly to work in Arnold Sommerfeld's institute in Munich and soon after in Copenhagen. Physicists in Europe were then on the verge of eliminating classical physics and providing the language essential to an understanding of the chemical bond. Pauling was there at exactly the right moment, with a mind perfectly well prepared to recognize the importance of what

the physicists were doing and then import it into chemistry. Frontiers are the seething fields of imagination, and where chemistry met physics was no exception. Here the book by Hager is much more comprehensible than that by the Goertzels and we see in it much more clearly how Pauling constructed his approach to the chemical bond and, later, to the contribution that he identified as his greatest, the concept of directed valence.

Meanwhile, the experimental weapons were being forged, as X-ray techniques were developed around the world, particularly by Lawrence Bragg at Manchester and later at Cambridge. Once again, Pauling had the right skills, a depth of

Roger Ressmeyer/SPL

intuition and a theoretical ability to cream the literature and solve the important questions of the day. The culmination of all this exposure to the finest experiments and the sharpest ideas was his proposal, worked out with the meticulous and complementary Robert Corey, of the structural motifs of proteins, the helices and sheets. Now, though, distractions set in, in the form of McCarthyism and its sniffing out of even the lightest pink from under the bed. Pauling, long a liberal, had been fanned by Ava Helen into action, and was clearly a target for the iniquitous domestic inquisition that marked the start of the Cold War. Too many of his acquaintances were bright red, and too vocal and skilled was his oratory in favour of a nuclear test ban, that it was inevitable that he would be a marked man. But he strenuously denied being a Communist, and although adamant that he would not reveal his political affiliations under duress, he nevertheless did deny membership of the party in a somewhat slippery manner when it appeared that his job was really on the line.

This kind of distraction, which is dealt with in detail in both biographies, inter-

fered with but by no means interrupted Pauling's scientific work. Yet now there was an air of haste. A Nobel prize that cemented his hubris in place, his failure to gain a passport to take him to Europe and an insufficient attention to detail led to his first major public mistake: the wrong structure for DNA. That seems to have been a turning point in his scientific evolution, for thereafter Pauling moved off into topics where hard science tends to melt.

In the seemingly twilight years of his work, we find Pauling's imagination and reading still as penetrating as ever, but now the vitamin C controversy takes centre stage and with it the ineradicable sense that in his seventh age Pauling became a crank. A particularly poignant episode is treated quite differently by the two books. Hager is firmly on Pauling's side on the issue of the summary dismissal from the Linus Pauling Institute of Science and Medicine at Menlo Park, California, of its president, Robinson. The Goertzel's present an account that is totally different (but recognizable), in which Pauling is the villain. To have these two views makes both biographies worthwhile. I am no better fitted to judge

where the exact truth lies than any one else who can read between the lines, but both accounts should be read.

Like his wife who predeceased him, Pauling also eventually succumbed to cancer; there is a suspicion that Ava Helen died partly on account of the megadoses of vitamin C that they both had taken. Yet his life was well lived and he has had a lasting influence on the way chemists think. We who are not in the immediate orbit of such a man benefit from his scientific insights, even if some of them (I have in mind the valence bond model) are no longer regarded as quite the bee's knees. We all perhaps benefit from his stand against witch-hunts and fallout.

Of these biographies, Hager's is plainly the more monumental and authoritative, but the touchstone test of the Robinson affair, and the almost total silence of certain members of Pauling's family in the book's pages, suggests that the author is concerned to paint in too rosy a hue. The Goertzel's book is much slighter, with spellings of names more phonetic than actual (which does not inspire confidence); yet it is written with greyer spectacles and deserves to be read alongside Hager's. Pauling, of course, speaks out clearly in his science, and it is also pleasing to have a brief well-presented anthology of his words set in their context and ranging over all his contributions. □

Peter Atkins is at Lincoln College, Oxford OX1 3DR, UK.

## Knowledge in the making

Edward J. Hackett

**Places of Inquiry: Research and Advanced Education in Modern Universities.** By Burton R. Clark. *University of California Press*: 1995. Pp. 284. \$40, £32.

IN "advanced nations" today, Burton Clark contends, "highly specialized and rapidly changing knowledge serves as a primary source of competence, intellectual energy, power, and wealth" for both individuals and organizations. People acquire such knowledge through organized study within "places of inquiry" dedicated to the generation and dissemination of new knowledge, so "one or more sectors of society must develop substantial capacity to generate new knowledge and rapidly disseminate it. Only nations willing to serve as saving remnants for received wisdom and past virtue... can ignore the modern need for a widening base of activities organized



A RARE blue iceberg in Antarctica, by the British photographer Cherry Alexander. The picture was the overall winning entry in the British Gas Wildlife Photographer of the Year Competition 1995, organized by BBC Wildlife Magazine and the Natural History Museum, London. The winning and commended photos from 22 countries are now reproduced in *Wildlife Photographer of the Year: Portfolio Five* (Fountain, £24.95).

around inquiry." Clark's task in this slim, well-structured volume is to examine the rise and social organization of the "research-teaching-study nexus" in a comparative institutional perspective, paying special attention to the role of outside organizations and forces in the integration of these dissimilar functions.

### National settings

Wilhelm von Humboldt married education to research in Germany during the early decades of the nineteenth century. Since then the Humboldtian ideal has taken distinctive forms in diverse national settings. Clark first locates such developments in the historical and institutional contexts of higher education systems in Germany, the United Kingdom, France, the United States and Japan, showing how each nation shaped its intellectual resources for advanced (postgraduate) education and research. Drawing on this case material, he then identifies forces that draw inquiry and advanced education closer together and those that drive them further apart. In his conclusion, he forcefully argues for the essential compatibility of research, teaching and advanced study and for the necessity of policies to secure the future of the research university. Each principal element of the book — the five chapters about national systems of research and advanced education, the two analytical chapters about forces of integration and disintegration, and the normative conclusion — makes a distinct contribution to scholarship.

The five nationally centred chapters offer fascinating overviews of the history and contemporary organization of higher education in 'advanced societies', with the essential character of each nation's system captured in a memorable phrase. In modern Germany, Humboldt's vision has been transfigured into the 'institute university', an organization that honours the ideal of integrated teaching, research and study in some institutions while abandoning it elsewhere. The UK variant is a 'collegiate university', which retains the intimate and expensive tutorial approach to undergraduate teaching while investing little in formal postgraduate education. Elite higher education in bureaucratic and centralized France is undertaken at the *grandes écoles*, in an arrangement Clark terms the 'academy university', with a sharp fall in quality and resources when one compares these intensely competitive institutions with the mass higher education available in the universities. The 'graduate department university' of the United States is a high-octane interpretation of Humboldt's vision, a "strikingly large, decentralized, diversified, competitive, and entrepreneurial" system where many large universities with formally organized graduate programmes compete for repu-

tation, resources and the best scholarly talent. At the 'applied universities' of Japan, engineering offers the best undergraduate education with a broad curriculum rich in fundamentals. But Japanese postgraduate education is weak, academic research is often tightly coupled to industrial interests, and graduate education in the humanities and social sciences scarcely exists.

The marriage of teaching and research is both fertile and fragile. Evidence for the fertility of the match may be found in the extraordinary productivity of the academic research enterprise, as indicated by international publication and citation data. Yet the marriage is uneasy, Clark warns us, for powerful centrifugal forces are at work in all the countries in his study. Teaching 'drifts' away from research with the growing burden of mass higher education and the rising demand for professional preparation. Research 'drifts' away from teaching with the widening gap between 'frontier knowledge' and 'teachable knowledge' and with the rising support for research by national and subnational governments, accompanied by rising demands for immediate, tangible, bankable results.

Countervailing centripetal forces at the national, university and departmental levels integrate teaching and research, offering the research university an anchor against both sorts of drift. Among the national-level 'enabling conditions' that draw teaching and research together is differentiation among universities, which allows a small sector to arise that integrates the two endeavours, while elsewhere in the post-secondary education system they remain separate. Intensified competition for resources, reputation and scholarly talent accompanies this segmentation. 'Formative conditions' at the university level — including the differentiation of an advanced 'graduate school' separate from the undergraduate college, the cross-subsidization of undergraduate and graduate activities and the institutional quest to construct and occupy a niche that bestows competitive advantage — are made possible by conditions at the national level and in turn create an organizational climate that promotes integration in academic departments. Departments integrate research, teaching and study through the joint action of intersecting groups: the research group, organized around a professor, some younger investigators and a coherent cluster of problems and methods, and the teaching group, represented by the department, which imposes admissions standards, curricular requirements, examination criteria and organized, collective scrutiny of young candidates' qualifications for advanced degrees.

Clark concludes by undermining and then transcending the belaboured 'ten-

sion' between teaching and research. As places of inquiry, universities endeavour to produce new knowledge, broadly defined to include ideas, empirical findings and critical perspectives, and then to convey that knowledge to others through publication, teaching and research experience. Pre-advanced or undergraduate education may rely more heavily on the communication of well-established knowledge, whereas advanced or graduate education deals in knowledge that is tacit, contingent, unsettled. But the spirit of inquiry pervades the university, transforming even professional education and enriching even the most basic courses. The prominence of inquiry in professional education is evidenced in the rise of such fields as 'management science', 'materials science' and 'basic research in manufacturing', replete with professional societies, conferences, scholarly journals and expectations of faculty to publish. Undergraduate institutions expect their faculty to publish, and undergraduate programmes are rapidly creating research experiences for students, experiential learning environments and 'studio' courses that integrate learning with inquiry. Tension between teaching and research is invoked chiefly as an ideology to justify cost-cutting measures on campuses, but the putative tension is both false and counterproductive. In its stead Clark proposes a "double helix... composed of intertwined parts of teaching and research, institutionally expressed as a teaching group and a research group...that holds the university against the tides of teaching drift and research drift. It counters the specific interests of government and industry that would otherwise place research in one isolated setting and teaching in another."

### Harsher criticism

*Places of Inquiry* is an admirably lucid book, historical, comparative and institutional in conception, synthetic in analysis and gently prescriptive in conclusion. One might wish for harsher criticisms of the social and political interests that motivate recent attacks on higher education and knowledge. One might also wish the author had examined places of inquiry in a less developed society, such as India, China or Brazil. Finally, one might ask Clark to apply his analysis and critical insights more directly to contemporary debates about higher education. But such small matters amount to a request for the author to write more and differently, when we should instead show appreciation for the graceful and enlightening volume in hand. □

Edward J. Hackett is in the Department of Science and Technology Studies, Rensselaer Polytechnic Institute, Troy, New York 12180-3590, USA.



## Deadly excuses

Kristie Macrakis

**Death and Deliverance: Euthanasia in Germany, 1900–1945.** By Michael Burleigh. Cambridge University Press: 1994. Pp. 400. £35, \$59.95 (hbk); £14.95, \$18.95 (pbk).

In addition to some 6 million Jews, 200,000 mentally ill or handicapped people were murdered by the Nazi regime. This year has marked the fiftieth anniversary of the Allied victory over this evil. Anniversaries often stimulate an introspective national examination of the past or at least call the attention of a wider audience to it. *Death and Deliverance* has come out at the end of a burst of literature on the general subject, a deluge which began on the eve of the fiftieth anniversary of the 1933 Nazi seizure of power.

At first glance, the volume appears to be yet another book on race hygiene, racial theory and Nazi doctors. But Michael Burleigh, a reader in international history at the London School of Economics and author of *Germany Turns Eastward* (Cambridge University Press, 1988), has instead added another brick to the edifice of knowledge about eugenics, medical science and the holocaust by focusing more exclusively on the euthanasia programme. Although this topic has been written about extensively in German, this is the first comprehensive study in English.

Euthanasia, or mercy killing (from the Greek *eu-thanatos* — easy death), usually means putting people to death, or withholding medical treatment from them, because of a painful or incurable disease. Voluntary euthanasia gives the patient a right to assisted death. In Nazi Germany this definition still held true, but added to the pool were many more cases of mentally ill or physically disabled people who were killed involuntarily.

During the 1920s the idea had already emerged in Germany of placing differing values on life: this notion was spread through Karl Binding and Alfred Hoche's influential book *Permission for the Destruction of Life Unworthy of Life*, first published in 1920. 'Lives unworthy of life' were found in psychiatric hospitals and were called "ballast lives" or "empty human husks".

Burleigh sees economic hardship and notions of racial fitness as the initial motives for German euthanasia in asylums during the 1920s and 1930s. This earlier programme escalated into a large-scale covert operation for mass murder during the Second World War. With diminishing economic resources and the necessities of war, the feeding of "useless

eaters" could not be justified. After a brief interpretative introduction, Burleigh for the most part lets the voices speak for themselves through extensive quotations of victims and perpetrators.

The Nazi euthanasia programme can be traced back to the fall of 1938 when a father named Knauer wrote to Adolf Hitler asking that his child, who was born blind, retarded and missing a leg and an arm, be allowed a "mercy death". His wish was granted. By May 1939 the child-murder operation had begun when Hitler asked Karl Brandt, his personal physician, to bring together an advisory committee to organize and prepare for the killing of retarded and deformed children. The programme was carried out in strict secrecy.

The state and the medical profession betrayed the trust of parents who decided to place their children in an asylum. The doctors had a standard list of concocted causes of sudden death such as brain oedema, appendicitis and pneumonia, explanations of which were then sent to the parents on institutional stationery. These were often unbelievable, especially, for example, when the child had had an appendix removed months before.

Early in the book we encounter Hermann Pfannmüller, the director of the Egfling-Haar asylum, who murdered children by slow starvation, through reducing their rations of food. Other methods such as lethal injection or poisoning might have attracted too much foreign hate propaganda.

The child euthanasia programme was the first major step leading to T-4, the massive covert adult euthanasia project to exterminate all German mental patients. The code name came from the administrative head office at Tiergartenstrasse 4 in Berlin. As elsewhere in the book, Burleigh provides us with detailed character and biographical sketches of the murderers. Although this is the book's strength it sometimes makes the reader lose track of the analytical framework and historical progression of the murderous project.

The final capstone in this rapid escalation of mass murder was the merging of the destruction of mental patients with the genocide of Jews at concentration camps in an operation with the code name 14f13. Burleigh points out that most of the doctors did not make a distinction between killing for racial, political or medical reasons.

Loathsome figures emerge from this narrative, and Burleigh unreservedly expresses his moral and personal repugnance for some of these doctors. He particularly dislikes Dr Friedrich Mennecke who was responsible for thousands of euthanasia killings and was one of the

T-4 doctors involved in 14f13. Much is known about the 14f13 operation and about Mennecke because he corresponded with his wife every day when they were separated (over the years he wrote some 2,000 letters). Burleigh quotes from one letter only: "this man was so thoroughly appalling — bullying, garrulousness and gluttony being the lesser vices that come to mind." The letters consist of "cloying infantilisms" ("Mein liebste Puttli-Mittlein..."), and Mennecke even describes his bowel movements: "Bim-bim-bim: 7 O'clock. Got up! A right good morning to you, mummikins! Kissy-wissies! You're dozing still, am I right? Now I'm off — to shave! Ahoy! 7:30 a.m. Ready, I've had a shit too...". These are rare glimpses into the mind of a psychopath (who was also a lieutenant SS colonel, which Burleigh does not mention).

With this example, Burleigh rejects the notion of Robert Jay Lifton (author of *The Nazi Doctors*) that the camp doctors lived double lives — killers at work and nice family men at home; rather, he argues that the doctors brought their families into the workplace. The notorious Professor Werner Heyde, for example, brought his wife and daughter to see selections at Dachau.

The 'selling' of murder was done through elaborate propaganda films, and here Burleigh's analysis is effective and graphic in its retelling of many of the infamous stories. The idea that death would lead to deliverance is a common theme in these films. Films about the mercy killing of terminally ill patients with diseases such as multiple sclerosis were used to persuade the audience of the worthiness of euthanasia while conflating it with the killing of mental patients.

The final chapter on the fate of the perpetrators after the war has almost become a standard element in books on the Nazi period. Here, as elsewhere, we witness how most of the worst murderers got off with light sentences and apologetically justified their work as "humane" with little remorse. Some even changed their names and found plum jobs in psychiatric hospitals.

Although the book is a detailed study of the extermination of the mentally ill, it makes little attempt to link this dress rehearsal to the murder of the Jews in concentration camps. Gas chambers equipped with showers were first used on the mentally ill. The technical apparatus for destroying other 'lives unworthy of living' was already there, as was the mind-set about certain categories of life. □

Kristie Macrakis is in the Department of History, Michigan State University, East Lansing, Michigan 48824-1036, USA.

# Making continental crust

Roberta L. Rudnick

**The continental crust has an andesitic bulk composition, which cannot have been produced by the basaltic magmatism that dominates sites of present-day crustal growth—at both convergent margins and within plates. These observations suggest that there may have been a different mode of continental crust generation in Archaean times, and may point to delamination of the lower crust as an important recycling process.**

ONE of the Earth's unique features when compared with other rocky planets in our Solar System is the presence of a chemically 'evolved' crust—the continental crust<sup>1</sup>. The continental crust covers ~40% of the Earth's surface area, and is defined laterally on the basis of the slope break on continental shelves<sup>2</sup> and vertically by a jump in seismic wave speeds to values greater than 7.6–8.0 km s<sup>-1</sup> (the Mohorovičić discontinuity). In contrast to the Earth's oceanic crust, which is predominantly thin, basaltic and young (<200 Myr), the continental crust is thick (20–60 km), with an average thickness of 39 km (ref. 3) and contains virtually every rock type known to geologists, including the oldest rocks found on Earth (presently the Acasta gneisses of Canada at 3.96 Gyr (ref. 4)). The composition of the continental crust is established by independent methods to be intermediate between basalt and rhyolite (in other words, andesitic), and it contains a significant fraction of incompatible trace elements—those that do not fit readily into crystal lattice sites and are therefore concentrated in melts. Thus the continental crust, although volumetrically insignificant (~0.57% of the mass of the mantle) is an important reservoir for the most incompatible elements.

Although the continents constitute one of the more accessible regions of the Earth, compared with the mantle, core and even the oceanic crust, there is considerable debate regarding how and when they formed and the processes responsible for their unique composition. Did the present mass of continents form very early in Earth's history (>4.0 Gyr ago), with past and present growth counterbalanced by recycling of crust into the mantle? If so, what were the main processes that formed the early crust and how do they compare with those operating today? And by what means are large masses of continental crust recycled into the convecting mantle? Alternatively, has the crust grown progressively, with important contributions having occurred after 2.5 Gyr, by processes roughly similar to those occurring now (island arc accretion, oceanic plateau accretion and basaltic underplating)?

Perhaps the greatest dilemma facing those interested in understanding how the continents formed is their composition. The continental crust is believed to be derived by partial melting of the mantle, and forms a complementary geochemical reservoir to the Earth's depleted mantle<sup>5</sup>. Yet most melting experiments carried out on mantle peridotite produce magmas that are considerably less evolved (for example, basalts and picrites) than the bulk continental crust (see, for example, refs 6 and 7).

To explain this paradox several models have been proposed. One suggests that the higher heat flow believed to be present earlier in Earth's history led to significant production of mantle-derived intermediate to silicic melts. These may have formed by melting of subducted oceanic crust<sup>8,9</sup>, by reaction between these melts and overlying peridotite<sup>10,11</sup>, or by shallow melting of wet, metasomatized peridotite<sup>12–14</sup>. The second<sup>15–18</sup> proposes that the crust grows by addition of basaltic magmas derived from the mantle (or by tectonic accretion of basaltic crust formed in island arcs or oceanic plateaus), which undergo intracrustal differentiation to form evolved melts and their complementary residues.

The former ascend buoyantly while the latter transform to a dense mineralogy (eclogite) in the deep crust and are lost to the convecting mantle during 'delamination' events that occur during continental collisions<sup>18</sup>. A third explanation is that chemical weathering of surface rocks leads to depletion in MgO, which exchanges with CaO during mid-ocean-ridge hydrothermal alteration and is ultimately returned to the mantle through subduction, thereby driving the crust towards more evolved compositions<sup>15</sup>. Finally, a fourth possible explanation is that ultramafic complements to the crust exist within the upper mantle as either eclogites or olivine-rich cumulates<sup>19,20</sup>, having high densities and acoustic velocities that are indistinguishable from those of residual peridotite<sup>21</sup>. This last suggestion is not borne out by studies of upper-mantle rocks brought to the surface as xenoliths in continental regions<sup>22,23</sup>, suggesting that if such voluminous cumulates did exist beneath the Moho, they have been largely lost from the lithosphere, perhaps by delamination.

Thus the first three explanations remain as possibilities and call for more detailed examination. These explanations are not mutually exclusive and it is possible that all have operated to some extent during the course of Earth history.

## When did the continents grow?

The timing of continental growth has been the topic of myriad papers over the past three decades and a definitive answer to this question has yet to be determined. Although space does not permit a thorough discourse on this issue, the following points are relevant here.

(1) The present age distribution of the crust, as determined from model age studies using the Sm–Nd isotope system (for example, refs 24–27) provides a minimum estimate of the crustal growth rate. These curves suggest that 35–60% of the present crustal mass formed in the Archaean, a considerably larger amount than the present areal extent of Archaean crust (14%, ref. 28). (2) The present age distribution of the crust does not constrain the maximum growth rate, owing to the occurrence of crustal recycling by means of sediment subduction, tectonic erosion at convergent margins and, perhaps, lower-crustal delamination<sup>29,30</sup>. Recent estimates of the flux of crustal material returned to the mantle at subduction zones are 1.3–1.8 km<sup>3</sup> yr<sup>-1</sup> (ref. 31), of which only a small portion returns to the crust in convergent margin magmas<sup>31,32</sup>. This volume is roughly comparable to current crustal production rates (for example, 1.65 km<sup>3</sup> yr<sup>-1</sup>, ref. 33) and it is uncertain how these rates may have changed in the past. Thus, anywhere between 40 to 100% of the present mass of the continents may have existed since the end of the Archaean.

Irrespective of when the continents attained their present mass, isotopic data demonstrate that continental formation has been occurring throughout Earth history. That is, continents are forming, but the net mass of the continents may or may not be changing. We can therefore use the cumulative mass curves determined from neodymium model ages to define approximately when the present mass of continents grew. It is the pro-

cesses responsible for this crust formation that are of interest here.

## Continental crust composition

Fundamental to understanding crustal growth is knowledge of the bulk crust composition. Thus it is appropriate in a review of crustal growth processes to begin by discussing the crust composition and evaluating how well it is known.

Estimates of bulk crust composition are derived from observations of the rock types present in the upper continental crust and models of lower-crust composition. The latter are derived from (1) proportions of rocks observed in outcrops of granulites (granulite terrains) that have equilibrated at lower-crustal pressures<sup>34, 36</sup>, (2) relating observed deep-crustal seismic velocities with rock types observed in both granulite xenoliths and terrains<sup>3, 37, 38</sup>, or (3) mixing mafic and silicic rock types in appropriate proportions to satisfy observed heat flow<sup>39</sup>. Thus current compositional estimates of the continental crust rely on diverse geochemical data bases (some specific to certain regions<sup>34, 36</sup>, others more global in nature<sup>37, 39</sup>); some make geophysical constraints cornerstones of the model<sup>37, 39</sup>, others do not<sup>34, 35</sup>.

Table 1 lists a current estimate of the bulk continental crust composition<sup>37</sup>. The major-element composition is fairly well defined, with various estimates falling within 30% of one another for most major elements (Fig. 1a). In contrast, there is variation of more than a factor of 2 in the estimates of K concentration, as well as those of other highly incompatible elements (for example, Cs, Rb, Th, U). The same is true for some compatible elements (for example, Ni, Fe, Fig. 1a). Such variations give rise to large uncertainties when calculating the amount of mantle that differentiated to form the crust (see, for example, refs 40, 41).

These uncertainties aside, there are some striking similarities in the crust compositional estimates reviewed above: they are all 'intermediate' in composition, with 57–64% SiO<sub>2</sub> and Mg# (defined as  $100 \times \text{molar Mg}/(\text{Mg} + \sum \text{Fe})$ ) between 50 and 56; they are all enriched in light rare-earth elements (REE), and have similar incompatible element ratios, with Nb and Ti depleted and Pb strongly enriched relative to the REE (Fig. 1b). Considering the diverse methods and observations used to create the crust compositional models, these similarities can be considered robust features of the continents with which any model of crustal growth must be in harmony. Below I explore how these features might be produced.

## Crustal growth processes

Continental growth is generally ascribed to two distinct plate tectonic settings: intraplate and convergent margins. Crustal growth at convergent margins, where lithospheric plates converge and basaltic oceanic crust is subducted into the mantle, may be accomplished by tectonic accretion of intra-oceanic island arcs or magmatic additions at continental arcs. Growth of continents in intraplate settings occurs in response to plume-related magmatism. This can happen within continents, such as in flood basalt provinces, or in ocean basins, resulting in anomalously thick oceanic crust (oceanic plateaus) that is difficult to subduct and is consequently accreted to continents during subduction-driven collisional events. Crustal growth in each of these settings is discussed in turn.

**The andesite model and its variants.** The andesitic composition of the continental crust, coupled with the spectacular andesite volcanism that occurs at convergent margin settings, led Taylor<sup>42, 43</sup> to propose that the continents form by accretion of island arcs of andesitic composition. This 'andesite model' of crustal growth appealed to uniformitarian sensibilities, in that processes we see occurring today could account for the formation of the continents. Subsequent investigations of continental crust and island arcs, however, have demonstrated the difficulties with this simple model. The andesite model of crust formation cannot account for the bulk-crust Cr and Ni contents (average

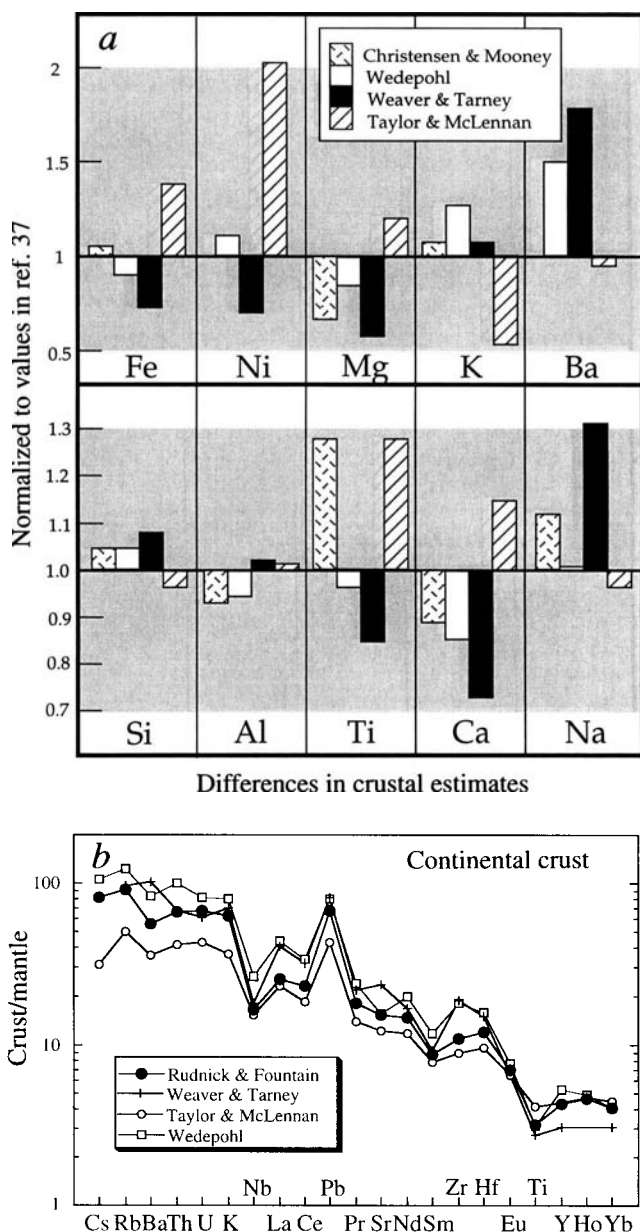


FIG. 1 a, Differences in elemental abundances between the various crust composition models. Values from Christensen & Mooney<sup>3</sup>, Weaver & Tarney<sup>34</sup>, Taylor & McLennan<sup>39</sup> and Wedepohl<sup>38</sup> are normalized to those of Rudnick and Fountain<sup>37</sup>. For the elements shown in the upper panel, variations between models are within a factor of two (shaded region), whereas for elements shown in the lower panel, variations are within 30% of the Rudnick and Fountain model (shaded region). b, Trace-element compositions of the continental crust. Data are normalized to primitive mantle<sup>73</sup> and plotted in order of their relative incompatibilities during mantle melting, as deduced from studies of oceanic basalts<sup>73</sup>. All models show high La/Nb, low Ce/Pb and similar Sr/Nd compared to undifferentiated mantle. These are considered robust features of continent composition.

andesites have abundances that are too low) nor its Th/U ratio<sup>39</sup>. Furthermore, a large portion of the continents probably formed during Archaean times (see above) and andesites are uncommon in Archaean volcanic sequences. This led Taylor and McLennan<sup>39</sup> to suggest that only post-Archaean crustal growth is accomplished by island arc accretion.

Perhaps most problematic for the andesite model, however, is that intra-oceanic island arcs are estimated to have basaltic, rather than andesitic bulk compositions<sup>15, 44, 46</sup>. This is consistent



with the observation that generally only basaltic or more magnesian magmas (for example, picrites or komatiites) are produced in high-pressure experimental melting studies of mantle peridotite. High-Mg# andesites and boninites that have equilibrated with mantle peridotite are thought to form under rather special  $P$ - $T$ - $X_{\text{H}_2\text{O}}$  ( $X_{\text{H}_2\text{O}}$  is molar proportion of water) conditions in modern arcs (that is, shallow, hot and wet)<sup>47,48</sup>, consistent with their relative rarity in arc settings. Thus accretion of modern island arcs produces basaltic crustal additions and cannot account for the intermediate composition of post-Archaean crust.

Andesites are more common in continental arc settings (such as the Andes, Cascades and central America) compared to intra-oceanic island arcs, presumably due to the effects of pre-existing continental crust on primary mantle-derived basaltic magmas (for example, ref. 16). But primary arc magmas in continental settings are estimated to have basaltic compositions, similar to those in intra-oceanic settings<sup>15</sup>.

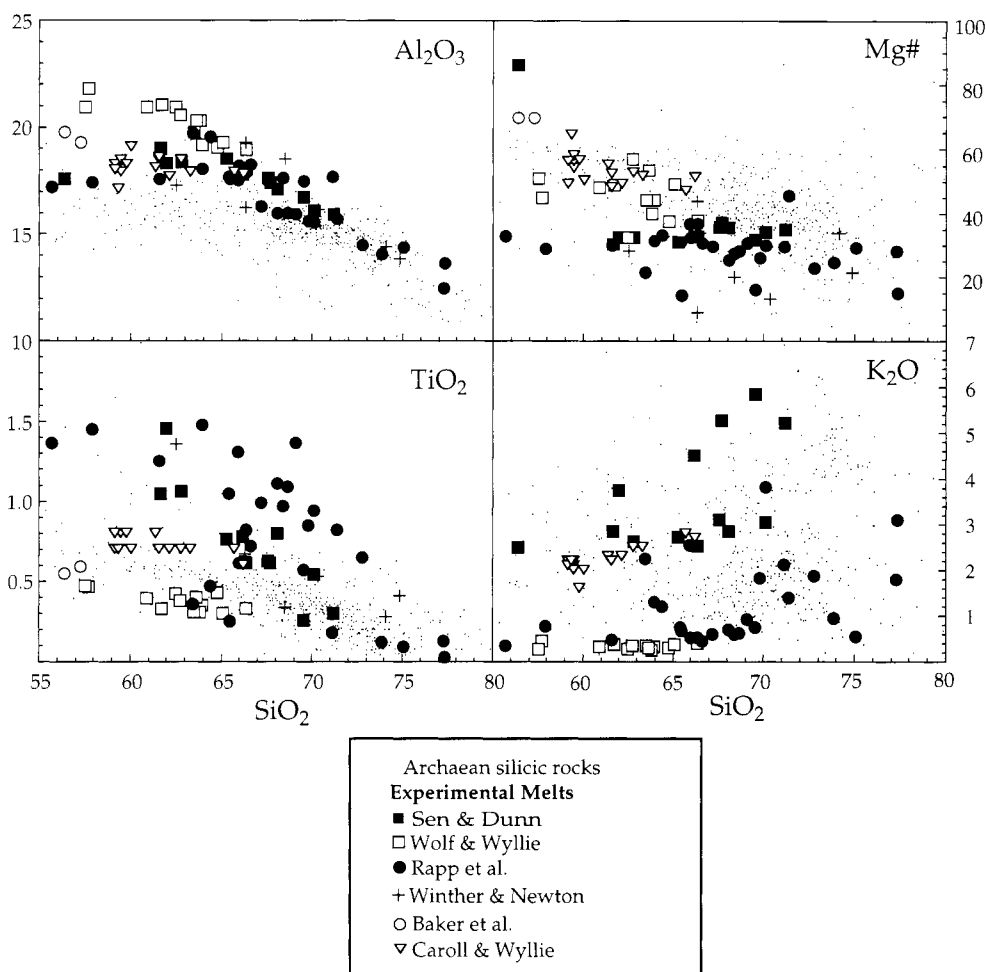
To overcome the above difficulties, several variations on the andesite model for crustal growth have been proposed. One of these suggests that during the Archaean, mantle temperatures may have been significantly higher than at present, leading to generation of mantle-derived intermediate to silicic melts. This may have occurred by melting of metasomatically enriched peridotite to produce intermediate melts<sup>7,9</sup> (the 'sanukitoids' of Shirey and Hanson<sup>12</sup>), which then differentiated to form silicic plutons, the so-called 'TTGs' (for tonalite-trondhjemite-granodiorite) that dominate in Archaean cratons. Alternatively, TTGs may have formed by direct melting of subducted oceanic crust, leaving behind a complementary eclogitic residue that is returned to the convecting mantle<sup>8,9</sup>. An important and distinctive feature

of TTGs is their trace-element compositions: they are enriched in incompatible trace elements and have strongly fractionated heavy REE patterns, suggesting that garnet played a role in their genesis (see ref. 49 and references therein).

Experimental studies have shown that production of andesite (or monzodiorites or quartz monzodiorites, as their intrusive equivalents are called) by melting of peridotite is possible under certain  $T$ - $X_{\text{H}_2\text{O}}$  conditions at uppermost mantle pressures. These include, at 1.0 GPa, temperatures in excess of 1,150 °C and magmatic water contents of 4–7 wt% (refs 20, 47, 48). A recent study produced andesite by 2% melting of anhydrous spinel lherzolite<sup>50</sup>, but a follow-up investigation has called these results into question (D. H. Green and T. Falloon, personal communication). Irrespective of the final resolution of this controversy, the very high viscosities of these melts<sup>50</sup> make them an unlikely means by which to generate the continents. Thus, volumetrically significant andesite production in the uppermost mantle appears to be limited to rather special hot and wet conditions that may have been more common in the Archaean than today<sup>14</sup>. Moreover, peridotite melting, whether hydrous or anhydrous, is not able to generate silicic melts with high Mg# nor the distinctive incompatible trace-element signatures of TTG. These features of TTGs, coupled with their primitive Nd-isotope compositions, require significant metasomatic enrichment of the peridotite source shortly before TTG genesis.

In contrast, production of tonalite and trondhjemite by partial melting of mafic rocks is well documented experimentally<sup>51–55</sup> and yields the requisite REE and Nb concentrations of TTGs, provided garnet and rutile are residual phases. In detail, the  $\text{TiO}_2$  and  $\text{Al}_2\text{O}_3$  contents of these experimentally produced melts are generally higher than those observed in Archaean TTGs at simi-

FIG. 2 Major-element variation diagrams comparing Archaean silicic rocks with melt compositions generated in experiments on amphibolite melting (Sen and Dunn<sup>54</sup>, Wolf and Wyllie<sup>55</sup>, Rapp et al.<sup>52</sup> and Winther and Newton<sup>53</sup>), peridotite melting (Baker et al.<sup>50</sup>) and mixed peridotite-tonalite melting (Carroll and Wyllie<sup>61</sup>).



lar  $\text{SiO}_2$  values (Fig. 2), whereas their Mg# are lower (especially at high  $\text{SiO}_2$  values). These discrepancies are probably due to the choice of experimental starting composition. Archaean basalts have, on average, higher MgO contents and Mg# and lower  $\text{Al}_2\text{O}_3$  and  $\text{TiO}_2$  contents than the basaltic compositions used in the amphibolite melting experiments (Fig. 3). Similarly, the relatively high Ni contents of some Archaean TTGs (see, for example, ref. 14) may reflect the Ni-rich character of Archaean basalts<sup>56</sup> coupled with a relatively low bulk partition coefficient for Ni during eclogite melting, where garnet, omphacite and rutile are likely to be residual phases. Unaltered Archaean basalts are expected to have low  $\text{K}_2\text{O}$  contents ( $\leq 0.1\%$ ), but if the chemical effects of sea-floor alteration in the Archaean were similar to those observed today (see, for example, ref. 57), then  $\text{K}_2\text{O}$  contents may be enhanced by an order of magnitude in altered Archaean oceanic basalts, thereby providing the  $\text{K}_2\text{O}$  necessary to form TTGs. Thus, despite some discrepancies between existing experimental melts and TTGs, choice of a starting composition that more closely matches altered Archaean basalts should yield silicic melts that are indistinguishable from Archaean TTGs.

What is not clear from the geochemistry of Archaean TTGs is at what depth they were generated; melting may have occurred within the lower continental crust, within subducted mafic crust or within metasomatized peridotitic mantle. If the first, the problem remains that voluminous mafic residues (garnet granulite or eclogite) are not present in either the lower crust or uppermost mantle of Archaean cratons<sup>58</sup>. The slab-melting model provides a convenient means of disposing of these residues, but it remains unclear what will happen to the silicic melts as they migrate through the overlying peridotitic wedge. Finally, the third model,

melting of metasomatized peridotite, can produce only high-Mg# andesites, yet these compositions are far outweighed by the more siliceous rocks of the Archaean crust (Fig. 2), some of which have equally high Mg#. These latter rocks cannot be produced from a metasomatized peridotite source.

As the trace-element characteristics of TTGs are the same as those predicted for a melt derived from subducted oceanic crust, it would seem that the two models of TTG genesis (slab versus peridotite melts) might profitably be combined into a model where TTGs are generated either by melting shallow mantle that has been hydrated and hybridized to pyroxenite by introduction of a silicic, slab-derived melt or by interaction of such a melt with overlying mantle wedge (see refs 10, 11 and references therein). It is possible that the volatile content of the hydrated crust served to buffer the degree of slab melting and that these melts interacted to variable extents on their way up through the mantle, thereby producing the distinctive and sub-parallel<sup>14</sup> trace-element patterns observed in TTGs (see Fig. 7 in ref. 10). Such models may also explain the presence of enstatite-rich peridotite found in some cratonic mantle xenolith suites<sup>59</sup> (for example, Kaapvaal craton<sup>60</sup>). Indeed, experimental investigations of peridotite-tonalite mixtures have produced silicic melts with high Mg#, similar to TTGs (Fig. 2) and reaction residues of orthopyroxene + garnet  $\pm$  clinopyroxene<sup>61</sup>.

Another variation on the andesite model has recently been proposed<sup>11</sup> in which andesites form in the mantle through interaction between rising basalts or silicic melts (the latter derived from subducted slabs) and peridotitic wall rock. In this model, young as well as ancient crust forms through direct addition of high-Mg# andesites. As the Mg# of the continental crust is significantly lower than that predicted for andesites which have

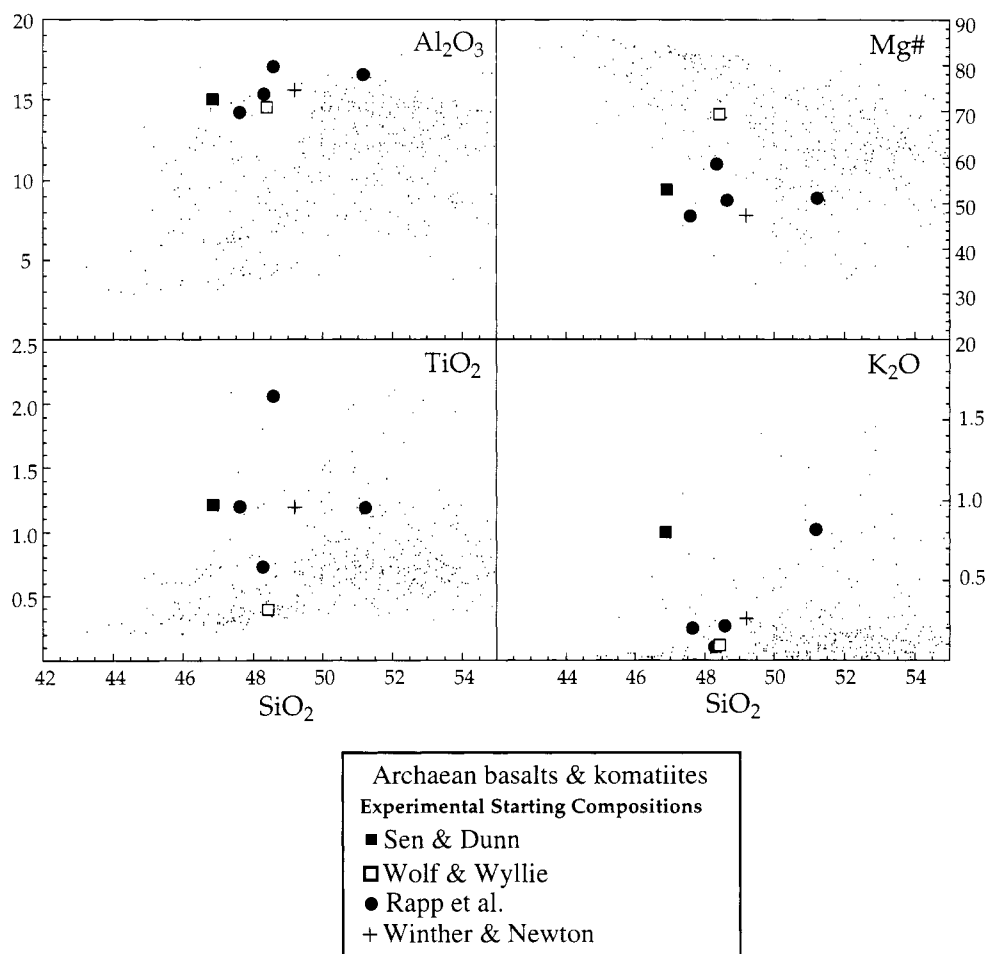


FIG. 3 Major-element variation diagrams comparing Archaean basalt and basaltic komatiite compositions with starting compositions used in amphibolite melting experiments (see Fig. 2). Data normalized to 100% anhydrous. In general, the Archaean lavas have lower  $\text{Al}_2\text{O}_3$  and  $\text{TiO}_2$  and higher Mg# than the compositions chosen for experimental investigation.  $\text{K}_2\text{O}$  contents for all but two of the basaltic starting compositions overlap values seen in Archaean basalts. It is possible that even higher  $\text{K}_2\text{O}$  contents were present in hydrothermally altered Archaean sea-floor basalts (see text).

equilibrated with peridotite (that is,  $Mg\# = 50\text{--}54$  for bulk crust, compared to  $>70$  for melts in equilibrium with peridotite), this model would require loss of high- $Mg\#$  cumulates from the crust or lithospheric mantle, as discussed above. Moreover, this model faces some of the same difficulties as the original andesite model for the generation of continental crust: namely, that intra-oceanic island arcs appear to have basaltic bulk compositions<sup>15,44,46,62</sup> and andesites are uncommon constituents of Archaean-aged crust<sup>63</sup>.

**Intraplate crustal growth.** During the Cenozoic, crustal growth at plate margins is estimated to outstrip growth at intraplate settings by a factor of 3 ( $1.1$  versus  $0.3\text{ km}^3\text{ yr}^{-1}$ , ref. 33), but it is not clear that this has always been the case. Large tracts of crust in west Africa<sup>64</sup> are interpreted to have grown by accretion of oceanic plateaus, considered to have formed by plume-related magmatism. In addition, oceanic plateaus make up about one-third of the accreted Wrangellia terrain in northwestern North America<sup>65,66</sup> and plume-associated magmatism has been estimated to contribute between 15% and >50% of some Archaean greenstone belts<sup>67,68</sup>. Intraplate magmatism may also contribute significant mass to continental crust via basaltic underplating in flood basalt provinces<sup>69,70</sup>. Plume-generated intraplate crustal growth may be episodic, volumetrically significant and ultimately related to mantle convection patterns<sup>71</sup>.

Can any limits be placed on the amount of crust generated in intraplate settings? As illustrated in Fig. 1 and discussed above, the continental crust has trace-element characteristics broadly similar to magmatic rocks generated in convergent margin settings. In detail, however, significant differences exist. Convergent margin magmas are characterized by variable, but generally high, Sr/Nd ratios compared to that of the primitive mantle (that is, 15; ref. 72). The Sr/Nd ratio of the continental crust overlaps these values, but lies at the lower end of the range (Fig. 4). That is, the continental crust has a lower Sr/Nd than is typical of convergent margin magmas. Likewise, the characteristic depletion in Nb relative to La in convergent margin magmas is also seen in the continental crust, but to a lesser degree (Fig. 4).

Convergent margin magmas with low Sr/Nd ratios also have negative Eu anomalies, owing to plagioclase fractionation. Thus, lower Sr/Nd in the continental crust compared with average arc volcanics may indicate that plagioclase-bearing cumulates, converted to plagioclase-free eclogitic mineralogies at the base of the thickened crust, have been lost from the continents by delamination of the lower crust (see below). The negative Eu anomaly seen in the bulk crust models of Gao *et al.*<sup>36</sup>, Wedepohl<sup>38</sup> and, to a lesser extent, Rudnick and Fountain<sup>37</sup> support this contention (Table 1). But, this process cannot explain the discrepancy between the La/Nb of the crust and convergent margin magmas because La/Nb, unlike Sr/Nd, does not vary significantly with depth in the continental crust<sup>37</sup> (that is, La/Nb is not significantly fractionated by intracrustal processes).

An alternative explanation of the trace-element characteristics of the crust is that they reflect growth through a mixture of convergent margin and intraplate magmatism. Intraplate magmas through time have Sr/Nd similar to the chondritic ratio (that is, 16) and La/Nb generally  $<1.0$  (refs 73, 74). Mixing relations between an intraplate basalt and possible convergent margin magmas are illustrated in Fig. 5. Between 10 and 35% of an intraplate component is required to explain the continental crust's La/Nb ratio.

It has been suggested on the basis of the light-REE enrichment of continental crust that its trace-element budget is dominated by the influence of small melt fractions from the mantle<sup>75</sup>. The dominantly arc-like trace-element characteristics of the continents highlighted above suggest that continent growth occurs mainly at convergent margins. In modern arcs, the role of fluids versus slab melts in generating the distinctive arc trace-element

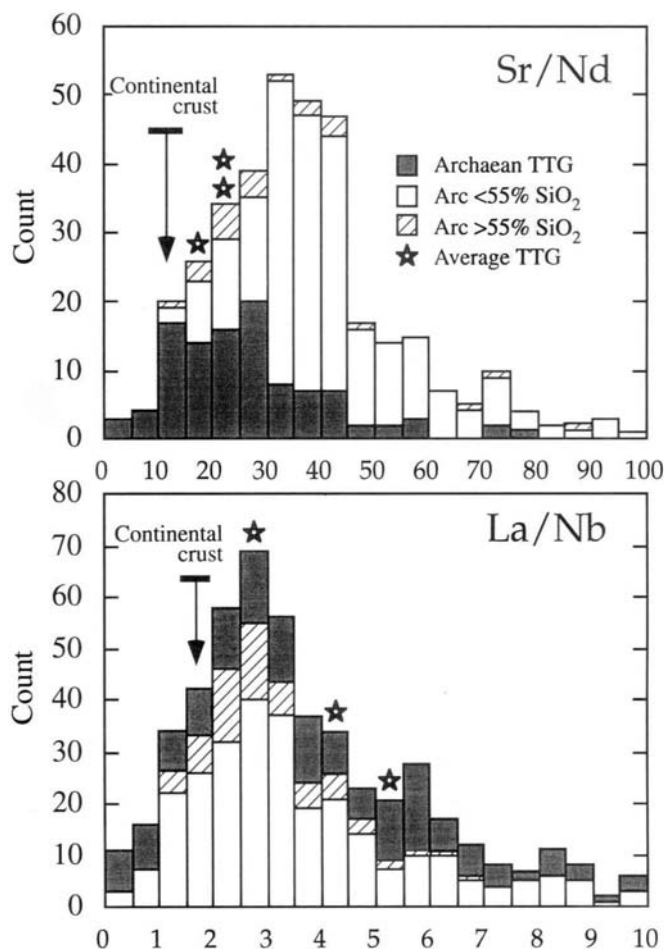


FIG. 4 Histograms of Sr/Nd and La/Nb ratios of arc magmas and Archaean TTGs (tonalites–trondhjemites–granodiorites) compared with bulk continental crust. To remove the effects of intracrustal differentiation on the Sr/Nd ratio, only rocks with  $0.90 < Eu/Eu^* < 1.10$  were included in the Sr/Nd plot ( $Eu/Eu^* = Eu_N / (Sm_N \times Gd_N)^{0.5}$ , where the subscripted N indicates the values are normalized to chondritic meteorites). Stars represent values of average TTG from refs 38, 49, 99.

signature is still debated, but most investigators acknowledge the importance of fluid transfer from slab to overlying mantle wedge (ref. 76 and references therein). The relative importance of these processes in the Archaean is more poorly constrained, but the models for TTG genesis discussed above suggest that slab melts of 10–20% may have been important. Thus, small-volume melts are not required to explain the crust's trace-element composition; fluid-transfer processes and slab melting may have played a more important role in generating the crust's composition.

In contrast to La/Nb, the Sr/Nd ratio of the continental crust is too low to be simply explained as a mixture of intraplate and convergent margin magmas (Fig. 5b). There are several possible explanations for this: (1) plagioclase-rich cumulates are under-represented in the lower-crustal xenolith populations that were used to model the composition of the lower crust, (2) Sr has been preferentially lost from the continents owing to carbonate precipitation in altered oceanic crust<sup>77</sup>, or (3) plagioclase-bearing cumulates have been lost from the base of the crust through delamination. Although the first explanation cannot be ruled out (because there are only ~70 determinations of Eu anomalies in mafic xenoliths currently available<sup>37</sup>), the other explanations are more appealing. Preferential Sr recycling is supported by the high Sr/Nd ratios observed in altered oceanic crust composites<sup>57</sup>, and delamination of a mafic to ultramafic layer from the base



TABLE 1 Bulk compositional estimate of the continental crust

|                                |      |        |        |
|--------------------------------|------|--------|--------|
| SiO <sub>2</sub>               | 59.1 | Ba     | 390    |
| TiO <sub>2</sub>               | 0.7  | La     | 18     |
| Al <sub>2</sub> O <sub>3</sub> | 15.8 | Ce     | 42     |
| FeO*                           | 6.6  | Pr     | 5.0    |
| MnO                            | 0.11 | Nd     | 20     |
| MgO                            | 4.4  | Sm     | 3.9    |
| CaO                            | 6.4  | Eu     | 1.2    |
| Na <sub>2</sub> O              | 3.2  | Gd     | 3.6    |
| K <sub>2</sub> O               | 1.9  | Tb     | 0.56   |
| P <sub>2</sub> O <sub>5</sub>  | 0.2  | Dy     | 3.5    |
|                                |      | Ho     | 0.76   |
| Mg#                            | 54.4 | Er     | 2.2    |
| Li                             | 11   | Yb     | 2.0    |
| Sc                             | 22   | Lu     | 0.33   |
| V                              | 131  | Hf     | 3.7    |
| Cr                             | 119  | Ta     | 1.1    |
| Co                             | 25   | Pb     | 12.6   |
| Ni                             | 51   | Th     | 5.6    |
| Cu                             | 24   | U      | 1.4    |
| Zn                             | 73   | K/U    | 10,100 |
| Ga                             | 16   | Sr/Nd  | 16     |
| Rb                             | 58   | La/Nb  | 1.5    |
| Sr                             | 325  | Ce/Pb  | 3.3    |
| Y                              | 20   | Nb/U   | 8.6    |
| Zr                             | 123  | Th/U   | 3.9    |
| Nb                             | 12   | Eu/Eu* | 0.96   |
| Cs                             | 2.6  |        |        |

These estimates are taken from ref. 37. Symbols used: Mg# =  $100 \times \text{molar Mg}/(\text{Mg} + \sum \text{Fe})$ ; FeO\*, total Fe as FeO; Eu\*, Eu value interpolated from REE pattern (see Fig. 4 legend).

of the crust could explain the major-element compositional paradox of the continental crust.

**Continental delamination.** Delamination of continental lithospheric mantle, owing to a density inversion between lithospheric and underlying asthenospheric mantle, was originally proposed to explain uplift of the Colorado plateau<sup>78</sup> and high-temperature metamorphism in the Himalayas<sup>79,80</sup>. Subsequently, delamination of the lower continental crust has been proposed to account for the intermediate composition of the continental crust<sup>17,18,81</sup>. A critical feature of lower-crustal delamination models is a density inversion between mafic or ultramafic lower crust and the underlying depleted peridotitic mantle. In these models, high-density lower crust may form as (1) ultramafic crystal cumulates from fractionating basaltic magmas at the base of the crust; these are rich in iron and hence denser ( $\rho = 3.4\text{--}3.6 \text{ g cm}^{-3}$ ) than refractory mantle composed of forsteritic olivine ( $\rho = 3.3 \text{ g cm}^{-3}$ , ref. 12) or (2) mafic lithologies (cumulates, restites or ponded melts) that convert to eclogite ( $\rho > 3.3 \text{ g cm}^{-3}$ ) at the base of an anomalously thick crust<sup>17,18</sup>.

Although the lower crust has yet to be included in numerical models of delamination, it appears that lithospheric thickening (such as occurs at sites of continental-scale collisions) is required to achieve delamination<sup>18,80</sup>. Thickening serves two purposes: (1) it promotes transformation of mafic lower crust into dense eclogite, and (2) it depresses cold, and presumably denser, lithospheric mantle into hotter and more buoyant asthenosphere, thereby providing the driving force for delamination. (Note, however, that compositional contrasts between fertile, asthenospheric peridotite and refractory peridotite, which may be characteristic of lithospheric mantle beneath arcs (for example, ref. 82) and Archaean cratons<sup>60,83</sup>, tend to counteract the density contrast caused by temperature differences). Previous or concurrent delamination of the underlying lithospheric mantle seems a prerequisite for ultimate loss of lower-crustal materials from the lithosphere<sup>18</sup>.

Although delamination of mafic to ultramafic rocks from the lower portions of the continents provides a means of explaining the non-basaltic composition of the crust, it is a difficult process

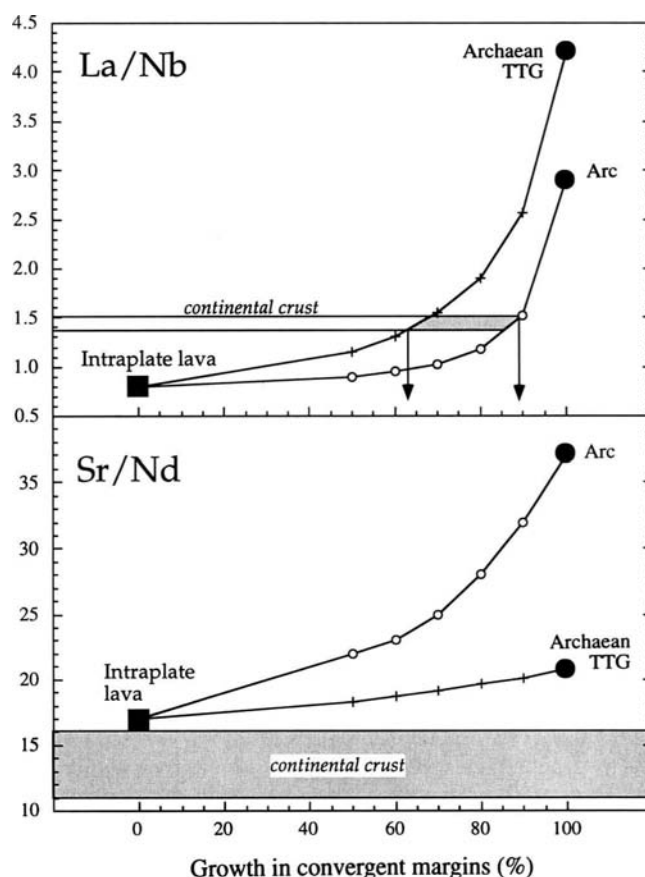


FIG. 5 Mixing relationships between convergent margin magmas (arc and Archaean TTG) and intraplate magma (average ocean island basalt from ref. 73); crosses and circles on curves represent 10% increments. Arc and Archaean TTG are median values from several hundred analyses compiled from the literature and plotted in Fig. 4. Continental crust bars represent range of ratios from estimates of refs 36–39. The model of Weaver and Tarney<sup>34</sup> is not considered here, as their choice of middle and lower crust is composed dominantly of Archaean TTG. Horizontal bars and shaded regions on both diagrams represent range of values for continental crust. Where these intersect the continental crust composition for La/Nb, the percentage of intraplate component can be read from the x-axis (marked by arrows). For Sr/Nd, the crust's ratio is lower than that of its potential 'building blocks', suggesting that the crust's Sr/Nd has been affected by other processes, such as preferential Sr loss during weathering or delamination of high-Sr/Nd material from the base of the crust. See text for full discussion.

to document. Some maintain that the geochemical characteristics of delaminated continental lithosphere are seen in plume-derived magmas from the ocean basins<sup>17,84</sup>, but these speculations are not evidence that the phenomenon occurs. Kay and Kay<sup>85</sup> and Kay *et al.*<sup>86</sup> suggest that sites of delamination in modern convergent margin settings may be recognized by rapid uplift and stress change (from compressional to extensional), low velocities and high attenuation of seismic waves (both features of hot mantle), and a temporal change in the composition of regional magmatism. The latter is regarded by Kay and Kay as providing the most geologically enduring evidence for delamination. They document changes in magmatism from the southern Puna plateau in Argentina that they attribute to delamination of the lower crust and upper mantle. Intrusions of granitic magmas, having geochemical evidence of being derived from a garnet-bearing source region (the thickened mafic lower crust in their model), are succeeded by basaltic magmas with intraplate chemistries (produced in response to incursion of hot asthenospheric mantle in the gap where the lithosphere used to be). It is possible, however, to produce these geochemical signatures in a variety

of ways. For example, a 'garnet signature' in granitic rocks may indicate partial melting of a mafic source at either the base of an anomalously thick lower crust, or within the basaltic layer of a subducted slab (as discussed above). Moreover, basalts having ocean-island type chemical characteristics, although uncommon, have been reported from a number of arcs (see discussion in ref. 87), which do not exhibit other evidence for lithospheric delamination. Given these uncertainties, recognizing delamination in older regions remains a difficult proposition.

It is desirable to determine whether delamination has occurred widely beneath the continents. The lower Sr/Nd ratio of the continental crust relative to its potential 'building blocks' (island arcs and oceanic plateaus) may argue for loss of rocks with high Sr/Nd from the base of the crust, as described above. Another way in which the importance of delamination might be tested is by evaluation of Os model ages for peridotite xenoliths derived from the continental lithospheric mantle. Unlike the other commonly utilized isotopic systems (Rb–Sr, Sm–Nd, U–Pb and Lu–Hf), the Re–Os isotopic system has the potential to date melt extraction, hence 'growth' and stabilization of lithospheric mantle<sup>88</sup>. If delamination occurs significantly after crustal growth, as may occur at sites of continental collision events, then lithospheric mantle in these regions should be younger than the overlying crust. To date, Os model ages are available for peridotites from only a few areas. In these areas the oldest Os model ages correspond well with the oldest ages observed in the overlying crust<sup>88–92</sup>. Thus, from the currently limited database, the Os model ages suggest either that (1) delamination has not occurred in these regions, or (2) if delamination occurred, it happened so close in time to original crustal growth that it is not detectable from the Os isotopic systematics (that is, within about 100 Myr). The latter may be true of continental regions that grow by accretion of island arcs or oceanic plateaus, where delamination occurs in response to their tectonic accretion to pre-existing crust shortly after their formation.

**Preferential Mg recycling.** The final process that may serve to shift the composition of the continental crust from basalt to andesite is preferential recycling of Mg to the mantle. Chemical weathering of igneous rocks results in loss of MgO, Na<sub>2</sub>O and CaO to solution; SiO<sub>2</sub>, Fe<sub>2</sub>O<sub>3</sub>, Al<sub>2</sub>O<sub>3</sub>, TiO<sub>2</sub> and K<sub>2</sub>O generally remain in the residues either as oxides or within or absorbed onto clay minerals<sup>93</sup>. Of the dissolved species, CaO is precipitated as carbonate and Na<sub>2</sub>O remains in solution or precipitates in evaporites. (Evaporites form the only known sedimentary sink for Na<sub>2</sub>O, but their volumes are so small (~5% of the sedimentary mass<sup>93</sup>) that they are not believed to influence the Na<sub>2</sub>O budget significantly<sup>94</sup>.) MgO may also precipitate in carbonates but a significant portion is fixed in oceanic crust through hydrothermal exchange with Ca at mid-ocean ridge spreading centres (ref. 95 and references therein). Thus, upon subduction of altered oceanic crust, Mg is preferentially recycled back to the mantle.

The recycling of Mg is occurring today and may have been even greater in the past if more extensive hydrothermal systems were present, perhaps due to greater ocean spreading-centre lengths. This process would serve to deplete the crust in MgO while other soluble elements are retained in, or returned to, the crust as sediments. However, Na<sub>2</sub>O is problematical because evaporites are uncommon and a good portion of the Na<sub>2</sub>O remains dissolved in the oceans. Attempts to recombine sedimentary rocks to form an 'average' igneous rock highlight the problems with Na<sub>2</sub>O (ref. 96).

Some limits on the effectiveness of chemical weathering in determining crust composition might be gained from the compositions of average sediments. As described above, fine-grained, clastic sedimentary rocks are notably depleted in Na<sub>2</sub>O and CaO. Sedimentary rocks of any sort comprise only a small fraction of the crust (~8% by mass<sup>39</sup>), so to shift the composition of continental crust by chemical weathering one must argue that many of the crystalline rocks in the crust contain a significant portion of formerly sedimentary material. Reprocessing of sedi-

ments within the crust produces crystalline rocks that are deficient in Na and probably Ca (depending on how much carbonate was present) relative to primary igneous rocks. S-type (sedimentary-type) granites<sup>97</sup> are an example of such a rock type. However, these granites are estimated to form less than 2% of all granitoids<sup>98</sup>. Indeed, the metaluminous composition of the bulk crust (Table 1) argues against the presence of a volumetrically significant weathered component within the crust.

In summary, preferential Mg recycling from the continents to the mantle due to chemical weathering has undoubtedly occurred for as long as subduction processes have been active on Earth. This process will drive the continents towards a more evolved bulk composition, but is unlikely to be the primary process by which the continental crust attained its present composition.

## Future directions

It is likely that each of the processes reviewed here has contributed to the production of the present continental crust. Future studies of the continents should develop tests of the relative importance of these processes in crustal growth and evolution. In the case of Archaean crustal growth and TTG genesis, experimental melting studies are needed for mafic compositions that better approximate Archaean basalts (lower TiO<sub>2</sub> and Al<sub>2</sub>O<sub>3</sub>, higher Mg# relative to modern basalts). Further experimental investigations into the interaction between hydrous silicic melts and refractory to fertile peridotite compositions would help in understanding the complex processes occurring in both ancient and modern subduction settings. Supporting evidence for crustal delamination can be sought in Os isotope studies; regions in which the lithospheric mantle is demonstrably younger than overlying crust might be *prima facie* evidence of delamination, although a coincidence in age of crust and mantle lithosphere is non-definitive. Numerical models of lower-crustal delamination that incorporate likely lithospheric temperature distributions at convergent margin settings and petrological constraints on phase transitions would aid in understanding the relative importance of the various driving forces behind delamination. Additional integrated geophysical, petrological and geochemical studies of potential present-day sites of lower-crustal delamination (such as the Tibetan plateau and the central Andes) would go far towards establishing the viability of this process. Finally, refinement of existing empirical estimates of continental crust composition are needed in order to place more quantitative constraints on models of crustal growth. □

R. L. Rudnick is at the Department of Earth and Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge, Massachusetts 02138, USA.

1. Taylor, S. R. *Tectonophysics* **161**, 147–156 (1989).
2. Cogley, J. G. *Rev. Geophys. Space Phys.* **22**, 101–122 (1984).
3. Christensen, N. I. & Mooney, W. D. *J. geophys. Res.* **100**, 9761–9788 (1995).
4. Bowring, S. A., Williams, I. S. & Compston, W. *Geology* **17**, 971–974 (1989).
5. Hofmann, A. W. *Earth planet. Sci. Lett.* **90**, 297–314 (1988).
6. Jaques, A. L. & Green, D. H. *Contr. Miner. Petrol.* **73**, 287–310 (1980).
7. Hirose, K. & Kushiro, I. *Earth planet. Sci. Lett.* **114**, 477–489 (1993).
8. Martin, H. *Geology* **14**, 753–756 (1986).
9. Drummond, M. S. & Defant, M. J. *J. geophys. Res.* **95**, 21503–21521 (1990).
10. Kelemen, P. B., Shimizu, N. & Dunn, T. *Earth planet. Sci. Lett.* **120**, 111–134 (1993).
11. Kelemen, P. B. *Contr. Miner. Petrol.* **120**, 1–19 (1995).
12. Shirey, S. B. & Hanson, G. N. *Nature* **310**, 222–225 (1984).
13. Stern, R. A. & Hanson, G. N. *J. Petrol.* **32**, 201–238 (1991).
14. Evans, O. C. & Hanson, G. N. in *Tectonic Evolution of Greenstone Belts* (eds Ashwal, L. & de Wit, M.) (Oxford Univ. Press, in the press).
15. Anderson, A. T. *J. geophys. Res.* **87**, 7047–7060 (1982).
16. Gromet, L. P. & Silver, L. T. *J. Petrol.* **28**, 75–125 (1987).
17. Arndt, N. T. & Goldstein, S. L. *Tectonophysics* **161**, 201–212 (1989).
18. Kay, R. W. & Kay, S. M. *Geol. Rundschau* **80**, 259–278 (1991).
19. Kay, S. M. & Kay, R. W. *Geology* **13**, 461–464 (1985).
20. Kushiro, I. *J. geophys. Res.* **95**, 15929–15939 (1990).
21. Griffin, W. L. & O'Reilly, S. Y. *Geology* **15**, 241–244 (1987).
22. Wilshire, H. W. et al. *Mafic and Ultramafic Xenoliths from Volcanic Rocks of the Western United States* (US Geological Survey, Washington DC, 1988).
23. Nixon, P. H. *Mantle Xenoliths* (Wiley, New York, 1987).
24. Patchett, P. J. & Arndt, N. T. *Earth planet. Sci. Lett.* **78**, 329–338 (1986).
25. McCulloch, M. T. in *Proterozoic Lithosphere Evolution* (ed. Kröner, A.) 115–130 (Am. Geophys. Union, Washington DC, 1987).
26. Jacobsen, S. B. *Geochim. cosmochim. Acta* **52**, 1341–1350 (1988).
27. DePaolo, D. J., Linn, A. M. & Schubert, G. *J. geophys. Res.* **96**, 2071–2088 (1991).
28. Goodwin, A. M. *Precambrian Geology* (Academic, London, 1991).

29. Armstrong, R. L. *Phil. Trans. R. Soc. Lond. A* **301**, 443–472 (1981).
30. Armstrong, R. L. *Aust. J. Earth Sci.* **38**, 613–630 (1991).
31. von Huene, R. & Scholl, D. W. *Rev. Geophys.* **29**, 279–316 (1991).
32. Plank, T. & Langmuir, C. H. *Nature* **362**, 739–743 (1993).
33. Reymer, A. & Schubert, G. *Tectonics* **3**, 63–77 (1984).
34. Weaver, B. L. & Tarney, J. *Nature* **310**, 575–577 (1984).
35. Shaw, D. M., Cramer, J. J., Higgins, M. D. & Truscott, M. G. in *The Nature of the Lower Continental Crust* (eds Dawson, J. B., Carswell, D. A., Hall, J. & Wedepohl, K. H.) 257–282 (Geol. Soc. London, London, 1986).
36. Gao, S. et al. *Geochim. cosmochim. Acta* **56**, 3933–3950 (1992).
37. Rudnick, R. L. & Fountain, D. M. *Rev. Geophys.* **33**, 267–309 (1995).
38. Wedepohl, H. *Geochim. cosmochim. Acta* **59**, 1217–1232 (1995).
39. Taylor, S. R. & McLennan, S. M. *The Continental Crust: its Composition and Evolution* (Blackwell, Oxford, 1985).
40. Allègre, C. J., Hart, S. R. & Minster, J.-F. *Earth planet. Sci. Lett.* **66**, 191–213 (1983).
41. Hofmann, A. W., Jochum, K. P., Seufert, M. & White, W. M. *Earth planet. Sci. Lett.* **79**, 33–45 (1986).
42. Taylor, S. R. *Tectonophysics* **4**, 17–34 (1966).
43. Taylor, S. R. in *Island Arcs Deep Sea Trenches and Back-arc Basins* (eds Talwani, M. & Pitman, W. C.) 325–336 (Am. Geophys. Union, Washington DC, 1977).
44. Arculus, R. J. *Tectonophysics* **75**, 113–133 (1981).
45. Eilam, R. M. & Hawkesworth, C. J. *Geology* **16**, 314–317 (1988).
46. Pearce, L. G., DeBari, S. M. & Sleep, N. H. *Earth planet. Sci. Lett.* **96**, 427–442 (1990).
47. Tatsumi, Y. *Earth planet. Sci. Lett.* **60**, 305–317 (1982).
48. Baker, M. B., Grove, T. L. & Price, R. *Contr. Miner. Petrol.* **118**, 111–129 (1994).
49. Martin, H. in *Archaean Crustal Evolution* (ed. Condie, K. C.) 205–259 (Elsevier, Amsterdam, 1994).
50. Baker, M. B., Hirschmann, M. M., Ghiorso, M. S. & Stolper, E. M. *Nature* **375**, 308–311 (1995).
51. Stern, C. R. & Wyllie, P. J. *Am. Miner.* **63**, 641–663 (1978).
52. Rapp, R. P., Watson, E. B. & Miller, C. F. *Precamb. Res.* **51**, 1–25 (1991).
53. Winther, K. T. & Newton, R. C. *Bull. Geol. Soc. Denmark* **39**, 213–228 (1991).
54. Sen, C. & Dunn, T. *Contr. Miner. Petrol.* **117**, 394–409 (1994).
55. Wolf, M. B. & Wyllie, P. J. *Contr. Miner. Petrol.* **115**, 369–383 (1994).
56. Arndt, N. T. *Tectonophysics* **187**, 411–419 (1991).
57. Staudigel, H., Davies, G. R., Hart, S. R., Marchant, K. M. & Smith, B. M. *Earth planet. Sci. Lett.* **130**, 169–185 (1995).
58. Ireland, T. R., Rudnick, R. L. & Spetsius, Z. *Earth planet. Sci. Lett.* **128**, 199–213 (1994).
59. Rudnick, R. L., McDonough, W. F. & Orpin, A. in *Kimberlites, Related Rocks and Mantle Xenoliths* Vol. 1 (eds Meyer, H. O. A. & Leonardos, O.) 336–353 (Proc. 5th Int. Kimb. Conf., C.P.R.M., Brasília, 1994).
60. Boyd, F. R. *Earth planet. Sci. Lett.* **96**, 15–26 (1989).
61. Carroll, M. R. & Wyllie, P. J. *J. Petrol.* **30**, 1351–1382 (1989).
62. Arculus, R. J. *Lithos* **33**, 189–208 (1994).
63. Condie, K. C. *Precamb. Res.* **32**, 261–278 (1986).
64. Abouchami, W., Boher, M., Michard, A. & Albareda, F. J. *geophys. Res.* **95**, 17605–17629 (1990).
65. Ben-Avraham, Z., Nur, A., Jones, D. & Cox, A. *Science* **213**, 47–54 (1981).
66. Richards, M. A., Jones, D. L., Duncan, R. A. & DePaolo, D. J. *Science* **254**, 263–267 (1991).
67. Hill, R. L., Campbell, I. H., Davies, G. F. & Griffiths, R. W. *Science* **256**, 186–193 (1992).
68. Vervoot, J. & White, W. M. *Eos (abstr.)* **73**, 617 (1992).
69. Cox, K. G. *J. Petrol.* **21**, 629–650 (1980).
70. Coffin, M. F. & Eldholm, O. *Rev. Geophys.* **32**, 1–36 (1994).
71. Stein, M. & Hofmann, A. W. *Nature* **372**, 63–68 (1994).
72. McDonough, W. F. *Earth planet. Sci. Lett.* **101**, 1–18 (1990).
73. Sun, S.-s. & McDonough, W. F. in *Magmatism in the Ocean Basins* (eds Saunders, A. D. & Norry, M. J.) 313–345 (Spec. Pub. 42, Geol. Soc. Lond., 1989).
74. McDonough, W. F. & Ireland, T. R. *Nature* **365**, 432–434 (1994).
75. O'Nions, R. K. & McKenzie, D. P. *Earth planet. Sci. Lett.* **90**, 449–456 (1988).
76. Stolper, E. & Newman, S. *Earth planet. Sci. Lett.* **121**, 293–325 (1994).
77. Goldstein, S. L. *Nature* **336**, 733–738 (1989).
78. Bird, P. *J. geophys. Res.* **84**, 7561–7571 (1979).
79. Bird, P. *J. geophys. Res.* **83**, 4975–4987 (1978).
80. Houseman, G. A., McKenzie, D. P. & Molnar, P. *J. geophys. Res.* **86**, 6115–6132 (1981).
81. Nelson, K. D. *Geophys. J. Int.* **105**, 25–35 (1991).
82. Jaques, A. L. & Chappell, B. W. *Contr. Miner. Petrol.* **75**, 55–70 (1980).
83. Jordan, T. H. in *The Mantle Sample: Inclusions in Kimberlites and other Volcanics* (eds Boyd, F. R. & Meyer, H. O. A.) 1–14 (Am. Geophys. Union, Washington DC, 1979).
84. McKenzie, D. & O'Nions, R. K. *Nature* **301**, 229–231 (1983).
85. Kay, R. W. & Kay, S. M. *Tectonophysics* **219**, 177–189 (1993).
86. Kay, S. M., Coira, B. & Viramonte, J. J. *geophys. Res.* **99**, 24323–24339 (1994).
87. Briggs, R. M. & McDonough, W. F. *J. Petrol.* **31**, 813–851 (1990).
88. Walker, R. J., Carlson, R. W., Shirey, S. B. & Boyd, F. R. *Geochim. cosmochim. Acta* **53**, 1583–1595 (1989).
89. Pearson, D. G. et al. *Mineralog. Mag.* **58A**, 703–704 (1994).
90. Carlson, R. W. & Irving, A. J. *Earth planet. Sci. Lett.* **126**, 457–472 (1994).
91. Pearson, D. G. et al. *Geochim. cosmochim. Acta* **59**, 959–977 (1995).
92. Pearson, D. G., Carlson, R. W., Shirey, S. B., Boyd, F. R. & Nixon, P. H. *Earth planet. Sci. Lett.* **134**, 341–357 (1995).
93. Garrels, R. M. & McKenzie, F. T. *Evolution of Sedimentary Rocks* (Norton, New York, 1971).
94. Goldschmidt, V. M. *Geochemistry* (Oxford Univ. Press, 1958).
95. Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans* (Princeton Univ. Press, Princeton, NJ, 1984).
96. Garrels, R. M. & McKenzie, F. T. *Evolution of Sedimentary Rocks* 242–243 (Norton, New York, 1971).
97. Chappell, B. W. & White, A. J. R. *Pacific Geol.* **8**, 173 (1974).
98. Miller, C. F. *Geology* **14**, 804–805 (1986).
99. Condie, K. C. *Chem. Geol.* **104**, 1–37 (1993).

ACKNOWLEDGEMENTS. We thank P. Hoffman and B. McDonough for providing helpful criticisms of the first draft. The final review has benefited from reviews by S. McLennan, P. Kelemen, R. Kay, J. Patchett, D. Abbott and C. Hawkesworth, and discussions with S. Eggins, W. McDonough and S. Jacobsen.

## ARTICLES

# Nucleation of microtubule assembly by a $\gamma$ -tubulin-containing ring complex

Yixian Zheng\*, Mei Lie Wong\*, Bruce Alberts\*<sup>†</sup> & Tim Mitchison<sup>†</sup>

\* Department of Biochemistry and Biophysics, † Department of Pharmacology, University of California at San Francisco, San Francisco, California 94143-0448, USA

**The highly conserved protein  $\gamma$ -tubulin is required for microtubule nucleation *in vivo*. When viewed in the electron microscope, a highly purified  $\gamma$ -tubulin complex from *Xenopus* consisting of at least seven different proteins is seen to have an open ring structure. This complex acts as an active microtubule-nucleating unit which can cap the minus ends of microtubules *in vitro*.**

ALL eukaryotes use the microtubule cytoskeleton for mitosis and cytoplasmic organization. Microtubules are long, hollow cylindrical structures (25 nm in diameter) formed from  $\alpha$ - and  $\beta$ -tubulin heterodimers<sup>1,3</sup>. Each microtubule is a polar structure with a plus and a minus end. The formation of a microtubule involves the addition of GTP-bound tubulin molecules to the growing end of the microtubule. *In vitro* studies show that a microtubule grows at a rate that is proportional to the concentration of free tubulin, and that at high tubulin concentrations the plus end grows approximately three times faster than the minus end<sup>4</sup>. *In vivo*, microtubule-organizing centres (MTOCs) nucleate microtubules and orient them so that each growing plus

end is distal to the MTOC. An MTOC catalyses nucleation at tubulin concentrations that are too low to promote spontaneous microtubule nucleation. In metazoa, the primary MTOC is the centrosome, consisting of a pair of centrioles surrounded by a cloud of electron-dense pericentriolar material (PCM), which is also known as the centrosome matrix. It is the PCM that is responsible for microtubule nucleation<sup>5</sup>. Microtubules nucleated from the PCM *in vivo* or *in vitro* have 13 protofilaments, whereas spontaneously assembled microtubules *in vitro* mostly have 14 protofilaments, with 13 and 15 protofilaments sometimes being found; it is therefore believed that MTOCs specify the exact structure of the microtubules that they nucleate<sup>6</sup>.

A very minor species of tubulin,  $\gamma$ -tubulin, has been identified as a new member of the tubulin superfamily<sup>7</sup>. Subsequently,  $\gamma$ -tubulin was found to be a highly conserved component of the

<sup>†</sup> Present address: National Academy of Sciences, 2101 Constitution Avenue NW, Washington, DC 20418, USA.



MTOCs from organisms ranging from fungi to humans<sup>8,9,12</sup>. Genetic analysis has demonstrated that  $\gamma$ -tubulin is required for both cytoplasmic and spindle microtubule formation in *Aspergillus nidulans*<sup>9</sup> and in *Schizosaccharomyces pombe*<sup>12</sup>. In addition, the use of antibodies to inhibit<sup>13</sup> or deplete<sup>14</sup>  $\gamma$ -tubulin blocks microtubule nucleation from the centrosome in higher eukaryotes. It has been suggested that  $\gamma$ -tubulin functions as the microtubule nucleator at the MTOC and that, by binding to the  $\beta$ -tubulin half of the tubulin molecule, it establishes the polarity of a microtubule—leaving the  $\alpha$ -tubulin half exposed at the plus end<sup>9</sup>. However, fluorescently labelled latex beads coupled to GTP bind specifically to the plus ends of microtubules<sup>15</sup>: as only  $\beta$ -tubulin can exchange its bound GTP, this result suggests that the  $\beta$ -tubulin half of the tubulin dimer is exposed at the plus end of a microtubule<sup>15</sup>.

To define the role of  $\gamma$ -tubulin in the nucleation of microtubule assembly further, it is necessary to study the biochemistry of the process. Such studies are difficult,  $\gamma$ -tubulin being a very minor protein in the cell, making it hard to purify, and genetic engineering cannot yet overproduce soluble functional  $\gamma$ -tubulin in either bacteria or baculovirus (ref. 16 and our unpublished observations). When an *in vitro* protein translation system was used to produce radioactively labelled  $\gamma$ -tubulin, the protein was found to be monomeric and to have the capacity to bind specifi-

cally to microtubules<sup>17</sup>. The endogenous *Drosophila* and *Xenopus*  $\gamma$ -tubulins, however, bind to microtubules as part of a large protein complex<sup>18,19</sup>. In *Drosophila*, several  $\gamma$ -tubulin complexes of different sizes were detected when cell extracts were analysed by sedimentation through sucrose gradients containing 500 mM NaCl<sup>18</sup>, but in analogous experiments using physiological salt conditions *Xenopus*  $\gamma$ -tubulin sediments as a single 25S particle<sup>19</sup>.

Here we report the purification of a *Xenopus*  $\gamma$ -tubulin complex and provide evidence that the purified complex can nucleate microtubule assembly by producing a new minus end of a microtubule.

### Purification of *Xenopus* complex

We characterized the state of  $\gamma$ -tubulin in *Xenopus* egg extracts by sedimentation in a sucrose gradient and determined its sedimentation coefficient as ~25S (data not shown) as reported previously<sup>19</sup> for the majority of *Xenopus*  $\gamma$ -tubulin. We then measured the Stokes radius of this apparently very large  $\gamma$ -tubulin complex as ~20 nm by gel filtration; the elution volume of  $\gamma$ -tubulin was determined by probing the column fractions for  $\gamma$ -tubulin on western blots (data not shown). On the basis of the sedimentation coefficient and the Stokes radius, we calculated the relative molecular mass for the  $\gamma$ -tubulin complex to be ~2,000K using the method described in ref. 20.

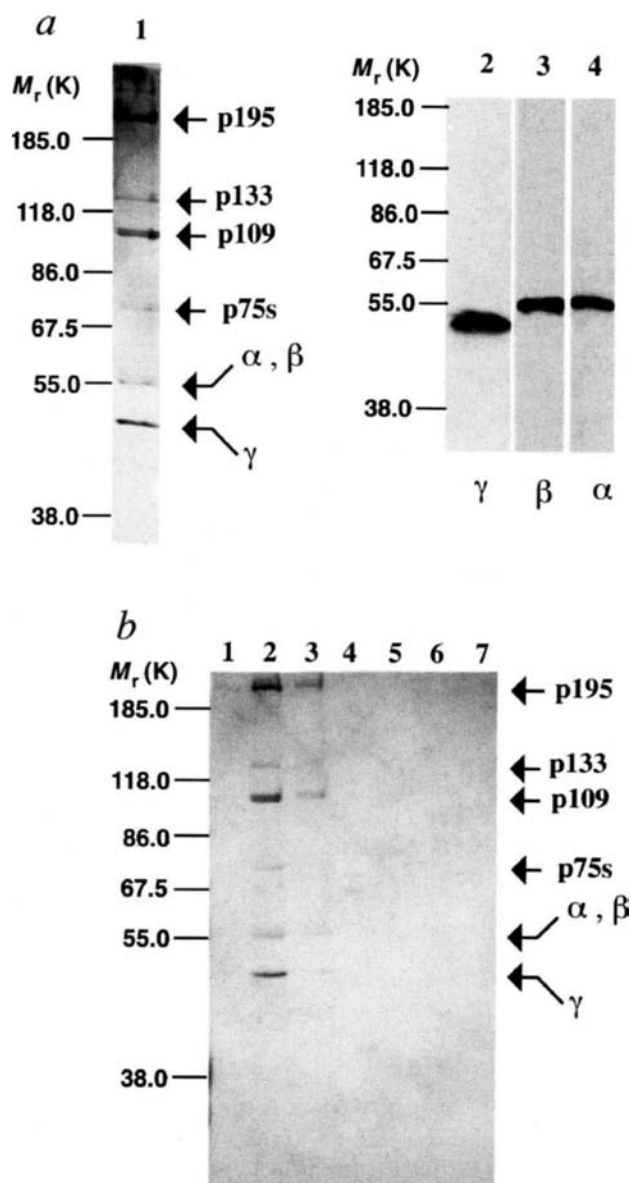


FIG. 1 Purification of the *Xenopus*  $\gamma$ -tubulin complex. *a*, Analysis by SDS-polyacrylamide gel electrophoresis of the  $\gamma$ -tubulin complex purified through the antibody-column step. Lanes: 1, gel stained with Coomassie blue to detect total protein; 2, western blot probed with an antibody (Xen C) against the C-terminal peptide AATRPDIYSWGTQDK of *Xenopus*  $\gamma$ -tubulin; 3, western blot probed with a monoclonal antibody (tub 2-28-33; Sigma) against  $\beta$ -tubulin; 4, western blot probed with a monoclonal antibody (DM1 $\alpha$ ; Sigma) against  $\alpha$ -tubulin. *b*, Coomassie-blue-stained gel of the proteins in *a* as they eluted from a subsequent SP-Sepharose Fast Flow column; lanes 1–7 are the salt step-gradient functions.

**METHODS.** We used antibodies raised against the C-terminal peptide of  $\gamma$ -tubulin because the C terminus should be exposed on the surface of the *Xenopus*  $\gamma$ -tubulin complex<sup>14</sup>. Polyclonal antibodies against the peptide (AATRPDIYSWGTQDK) were raised and purified as described<sup>33</sup>. An antibody column was made by coupling 1.5–2 mg of affinity-purified antibodies to 1 ml protein A-agarose beads (GIBCO BRL) as described<sup>33</sup>. All of the following purification steps were done at 4 °C. A 25–30-ml concentrated crude extract made from CSF-arrested *Xenopus* eggs<sup>34</sup> was clarified by centrifugation and then precipitated by addition of a saturated ammonium sulphate solution. The pellet from the 15–25% ammonium sulphate cut contained ~80–90% of the total  $\gamma$ -tubulin (typical yield ~8  $\mu$ g); it was dissolved in 10 ml HEPES-100 buffer (50 mM sodium-HEPES, pH 8.0, 1 mM EGTA, 1 mM MgCl<sub>2</sub>, 100 mM NaCl) containing 0.1 mM GTP, and clarified by centrifugation. The supernatant was further purified on a Sephacryl S-300 HR (Pharmacia) sizing column using the same buffer (the bed of 350 ml final volume was packed in an XK 26/70 column housing; Pharmacia). 5-ml fractions were collected at a flow rate of 3 ml min<sup>-1</sup>. This step gives about 70% recovery of  $\gamma$ -tubulin (~5  $\mu$ g) by quantified western blotting. The  $\gamma$ -tubulin-containing fractions (40 ml) were then passed through the 1-ml antibody column at a flow rate of 25 ml h<sup>-1</sup>. After washing the column in HEPES 100 buffer containing 0.1 mM GTP,  $\gamma$ -tubulin complexes were eluted with 0.4 mg ml<sup>-1</sup> peptide (AATRPDIYSWGTQDK) in the same buffer at 25 ml h<sup>-1</sup>. Pooled eluate fractions (~2 ml), which contained about 50% of the initially loaded  $\gamma$ -tubulin (yield ~2  $\mu$ g  $\gamma$ -tubulin), were loaded onto a 200- $\mu$ l SP-Sepharose Fast-Flow (Pharmacia) column packed in a micropipette tip and run at about 1.2 ml h<sup>-1</sup> by gravity. The peptide did not bind to the column. Bound  $\gamma$ -tubulin complex was eluted with a 1.4-ml salt gradient (100–700 mM NaCl in HEPES 100 buffer containing 0.1 mM GTP). Recovery was ~25–40%; final yield of total protein was 1–3  $\mu$ g ( $\gamma$ -tubulin final yield, 200–800 ng) in 400  $\mu$ l as quantified by densitometry using BSA or tubulin standards. We found that bovine tubulin consistently binds more Coomassie blue than BSA, but the increase in intensity is within 50%. The estimated stoichiometry of  $\gamma$ TuRC from two separate preparations was: preparation 1, for every 10 molecules of  $\gamma$ -tubulin, there are 1 of  $\alpha/\beta$  tubulin, 0.8 of p75, 2.6 of p109, 1.2 of p133 and 4 of p195; preparation 2, for every 10  $\gamma$ -tubulin molecules, there are 1.7  $\alpha/\beta$  tubulin, 2.3 p75, 2.8 p109, 0.6 p133 and 2.6 p195 molecules (estimated error,  $\pm$ 30%).

We estimated the amount of  $\gamma$ -tubulin in concentrated *Xenopus* egg extract by first immunodepleting  $\gamma$ -tubulin from the extract. The immunoprecipitated  $\gamma$ -tubulin was then fractionated by SDS-polyacrylamide gel electrophoresis run with known amounts of pure bovine serum albumin and bovine tubulin as standards. After staining the gel with Coomassie blue, we quantified the band intensity of  $\gamma$ -tubulin and the standards by densitometric scanning, finding that  $\gamma$ -tubulin makes up only  $\sim 0.001\%$  of the total extract protein (data not shown).

To simplify the purification of the  $\gamma$ -tubulin complex from *Xenopus*, we used antibody-affinity chromatography as part of the procedure (Fig. 1 legend). At least six other proteins copurified with  $\gamma$ -tubulin at this step (Fig. 1a) and through a subsequent cation-exchange column (Fig. 1b). In western blot analysis,  $\gamma$ -tubulin is the smallest protein in the set of copurifying polypeptides (Fig. 1a, lanes 1 and 2). The 55K band was identified as unresolved  $\alpha$ - and  $\beta$ -tubulins (Fig. 1a, lanes 3 and 4). Densitometry showed that  $\alpha$ - and  $\beta$ -tubulins were present in equal amounts (data not shown), suggesting that they form a normal  $\alpha$ - and  $\beta$ -tubulin heterodimer. Other copurifying proteins (arrows in Fig. 1a) are designated by their apparent molecular weights as p75, p109, p133 and p195. p109 is present as two bands on the gel, whereas p75 consists of three or four closely spaced bands. Because all the proteins shown in Fig. 1b cosediment on a sucrose gradient at  $\sim 25S$  (data not shown), we believe that they all belong to the same  $\gamma$ -tubulin complex.

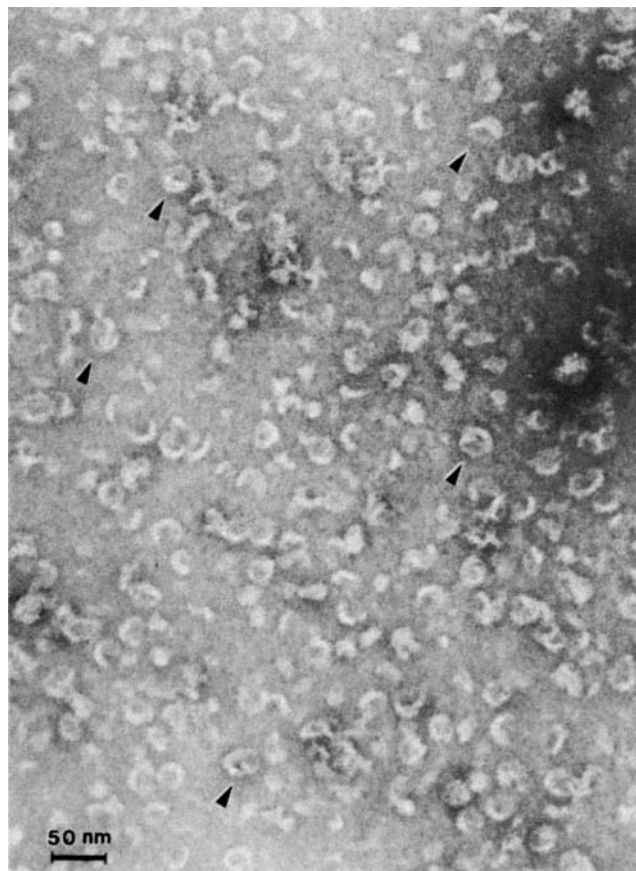


FIG. 2 Electron microscopy of a field of negatively stained  $\gamma$ -tubulin complexes; arrowheads point to rings that appear as left-handed helices.

**METHODS.** After elution of the  $\gamma$ -tubulin complex from SP-Sepharose (Fig. 1 legend), 5  $\mu$ l of each fraction was applied to 200-mesh carbon-coated grids. After 2–5 min, grids were rinsed in water, stained with uranyl acetate in 50% ethanol for 30 s and examined in a Philips EM400 microscope.

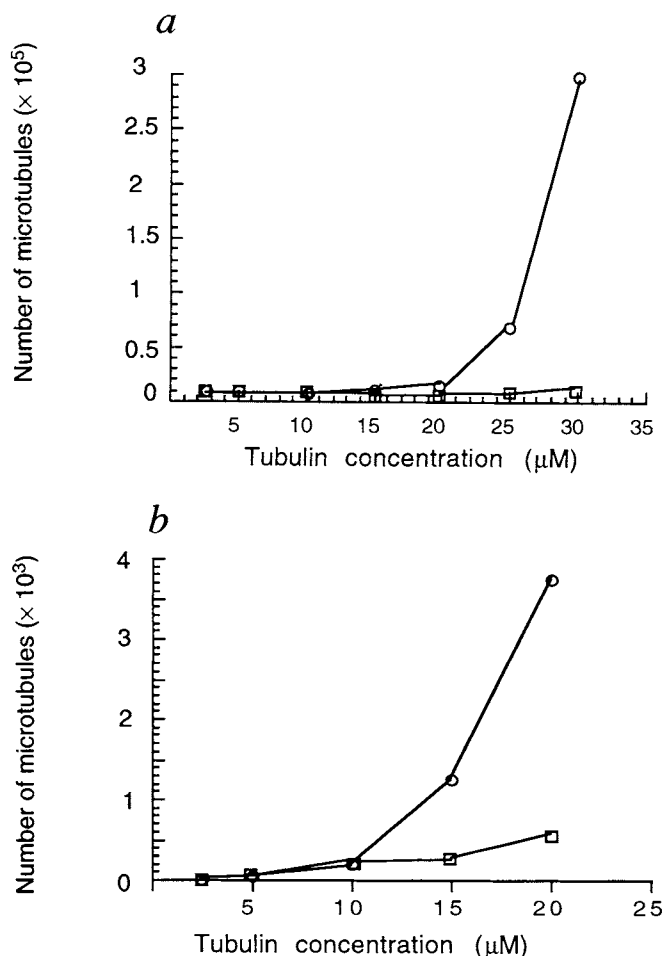


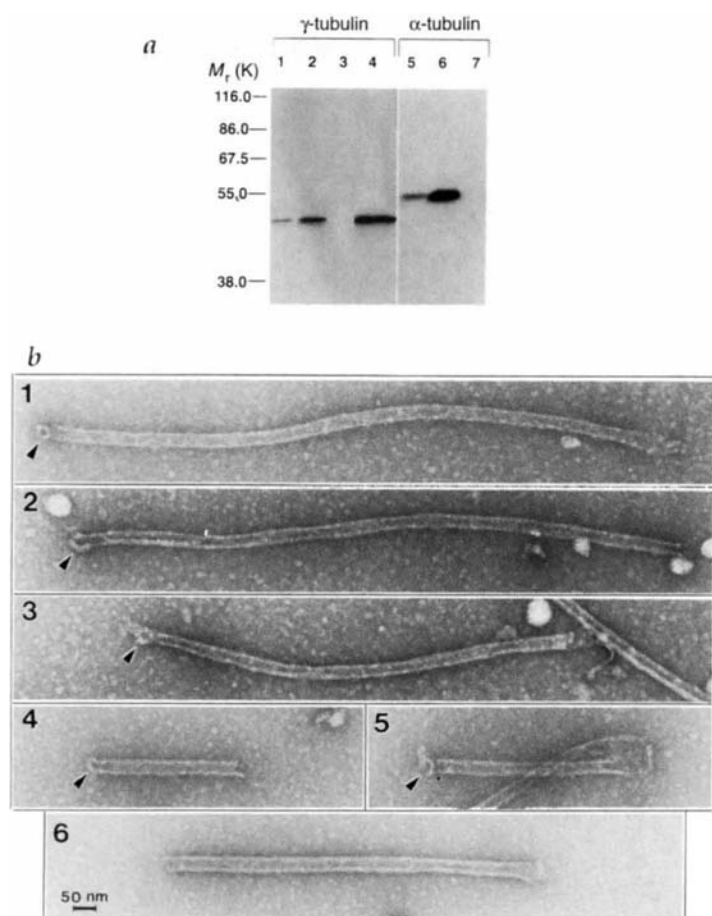
FIG. 3 Nucleation of microtubules by purified  $\gamma$ TuRC. Microtubules were grown with the indicated concentrations of tubulin in the presence of either  $\gamma$ TuRC (circles) or buffer (squares) and the number of microtubules observed after a fixed time was plotted. a, Complete plot; b, expanded plot of a to show the point at which  $\gamma$ TuRC starts to nucleate more microtubules than the control.

**METHODS.** Rhodamine-labelled and unlabelled bovine tubulin were prepared as described<sup>35</sup>. Tubulin concentration was determined using a calculated extinction coefficient of  $\epsilon_{280} = 88,500$ ; as bovine tubulin sequences were not available, we used pig  $\alpha/\beta$  tubulin sequences for this calculation<sup>36</sup>. The contribution of GTP to the  $A_{280}$  was subtracted from the total absorption. Microtubules were nucleated at the indicated tubulin concentrations in BRB80 buffer (80 mM K-PIPES, pH 6.8, 1 mM  $MgCl_2$ , 1 mM EGTA, 1 mM GTP) in the presence of either 4.3  $\mu$ g  $ml^{-1}$  purified  $\gamma$ TuRC or a buffer control. Rhodamine-labelled tubulin was used (1 labelled to 5 unlabelled tubulin molecules) to facilitate observation. Reactions were incubated at 37  $^{\circ}C$  for 5 min and then fixed with 1% glutaraldehyde. Microtubules in 50 random fields were counted at a magnification of  $\times 600$ , so these numbers represent only a very small fraction of the total microtubules nucleated.

To estimate the stoichiometry of the proteins in the purified  $\gamma$ -tubulin complex, we determined the relative amount of each band in the gel using bovine serum albumin and tubulin as standards (Fig. 1 legend). For every ten  $\gamma$ -tubulin molecules, we estimated that there are two or three p195, half or one p133, four or five p109, one or two p75, and one or two tubulin molecules. As the relative molecular mass of the entire complex is  $\sim 2,000K$ , we estimate that there are 10–13  $\gamma$ -tubulin molecules per molecule of  $\gamma$ -tubulin complex.

### Structure of the purified complex

When the purified complex is negatively stained and examined by electron microscopy, a ring structure is observed that is 25–28 nm in diameter (Fig. 2). Many of these rings appear to be



**FIG. 4** Assays of binding of  $\gamma$ TuRC to microtubules. **a**,  $\gamma$ TuRC binds to taxol-stabilized microtubules: purified  $\gamma$ TuRC was incubated with preformed microtubules (or with a buffer control) and then centrifuged through a glycerol cushion; western blots of the pellets are shown. Lanes: 1 and 2, microtubules added, with twice as many microtubules in lane 2 as in lane 1; 3, buffer control without microtubules; 4, total amount of  $\gamma$ TuRC used in each experiment; 5, 6 and 7, same samples as in lanes 1, 2 and 3, respectively; lanes 1–4 were probed with an antibody against  $\gamma$ -tubulin, Xen C, and lanes 5–7 were probed with a monoclonal antibody against  $\alpha$ -tubulin, DM1 $\alpha$ . **b**, Electron microscopy of microtubules nucleated in the presence or absence of the purified  $\gamma$ TuRC. Panels 1–5, selected microtubules nucleated in the presence of the  $\gamma$ TuRC; arrowheads point to a ring structure at one end of each microtubule. Panel 6, a typical microtubule nucleated in the absence of the  $\gamma$ TuRC.

**METHODS.** **a**, Microtubules were produced by incubation of 15  $\mu$ M tubulin in 190  $\mu$ l BRB80 buffer for 20 min at 37 °C. Then 200  $\mu$ M taxol was added and the incubation continued for another 10 min. Microtubule aggregates were then removed by centrifugation at 12,000 g for 5 min. Either 16 or 32  $\mu$ l microtubule-containing supernatant, or BRB80 buffer alone, as a control, was incubated with 5  $\mu$ g ml<sup>-1</sup> (final concentration) of purified  $\gamma$ TuRC in 50  $\mu$ l. After 10 min at room temperature, each reaction was spun at 50,000 r.p.m. for 10 min through a glycerol cushion (30% v/v glycerol in BRB80) in a TLA100 rotor (Beckman). The pellet was resuspended in SDS-sample buffer and analysed by western blotting after SDS-PAGE. **b**, Microtubules were produced by incubating 25  $\mu$ M tubulin in BRB80 at 37 °C for 40 s in the presence or absence of 6.4  $\mu$ g ml<sup>-1</sup> (final concentration) of purified  $\gamma$ TuRC. After fixation with 1% glutaraldehyde, microtubules were spun through a glycerol cushion (25% v/v glycerol in BRB80) onto carbon-coated 400-mesh EM grids. Grids were processed for negative staining and examined as described for Fig. 3. Microtubules showing a ring structure (arrowheads), selected from the reactions containing  $\gamma$ TuRC, are displayed in panels 1–5; panel 6 shows a control microtubule.

open, with the two ends overlapping each other. As the rings sit on the surface in all orientations, we were able to measure both the width and height dimensions of the non-overlapping region of the ring (the ring wall), which are both  $\sim$ 10 nm. Thus, the wall of a ring is roughly cylindrical, and it is both wider than the wall of a microtubule ( $\sim$ 4 nm wide) and higher than an  $\alpha/\beta$  tubulin dimer ( $\sim$ 8 nm high). Some of the rings clearly have the form of a left-handed helix, with more than one but less than two turns (arrowheads). Most structures that we see in these micrographs have the same ring structure, providing further evidence for the purity of the  $\gamma$ -tubulin complex. We will refer to this complex as the  $\gamma$ -tubulin ring complex, abbreviated as  $\gamma$ TuRC.

### $\gamma$ TuRC nucleates microtubules *in vitro*

To study the microtubule-nucleating ability of  $\gamma$ TuRC, we incubated the same amount of  $\gamma$ TuRC with increasing concentrations of purified bovine tubulin and measured the number of microtubules polymerized after 5 min at 37 °C. As controls, we substituted buffer for the  $\gamma$ TuRC in parallel reactions. As shown in Fig. 3a, there is very little microtubule formation at a tubulin concentration of  $\leq$ 10  $\mu$ M in the presence or absence of  $\gamma$ TuRC. But at 15  $\mu$ M tubulin, more microtubules are formed in the presence of  $\gamma$ TuRC than in the buffer control (Fig. 3b). The difference becomes striking at 30  $\mu$ M tubulin: although more microtubules now nucleate spontaneously, these make up only 1.5% of the total microtubules formed in the presence of  $\gamma$ TuRC (Fig. 3a).

In these experiments, we used  $\sim 1.5 \times 10^{10}$   $\gamma$ TuRC molecules per reaction (as estimated from the protein on a Coomassie-blue-stained gel). At the highest tubulin concentration tested (30  $\mu$ M), about  $2 \times 10^{10}$  microtubules were nucleated in the presence of the  $\gamma$ TuRC, versus  $3 \times 10^8$  microtubules in its absence. Thus, the number of microtubules nucleated in the presence of

the  $\gamma$ TuRC is about the same as the number of the  $\gamma$ TuRCs. This suggests that each  $\gamma$ TuRC might nucleate one microtubule through the type of end-binding mechanism previously suggested by Oakley *et al.*<sup>9</sup>

### Binding to one end of a microtubule

To determine whether the  $\gamma$ TuRC binds to microtubules, we incubated the purified  $\gamma$ TuRC with performed, taxol-stabilized microtubules and then centrifuged the solution. Antibody tests revealed that  $\gamma$ TuRC sediments with microtubules (Fig. 4a, lane 1) and that more  $\gamma$ TuRC sediments into the pellet when the amount of microtubules is increased (Fig. 4a, lane 2). As no  $\gamma$ TuRC was found in the pellet in the absence of microtubules (Fig. 4a, lane 3), we conclude that the  $\gamma$ TuRC can bind to preformed microtubules.

Using electron microscopy to localize  $\gamma$ TuRC on the microtubules grown in its presence, we observed that about 20% (11 of 56) of the  $\gamma$ TuRC-nucleated microtubules had a distinctive structure at one of their ends; moreover, over half of these structures (7 of 11) were clearly rings (Fig. 4b, panels 1–5). No such structures were found at the ends of spontaneously nucleated microtubules (Fig. 4b, panel 6). These results show that  $\gamma$ TuRC can bind to microtubules and that, at least for a fraction of the  $\gamma$ TuRC, the binding site is at one end of a microtubule.

### Blocking growth at the minus end

We would expect the binding of  $\gamma$ TuRC (Fig. 4b) to block (or at least retard) the addition of tubulin to whichever microtubule end is bound. If the  $\gamma$ TuRC nucleates microtubule growth in the centrosome, it might be expected to bind to the minus end of a microtubule. To find out which end is actually bound, we took advantage of the different rates of growth at each of the two ends of a microtubule: at a given concentration of free tubulin, the plus end will grow approximately three times faster than the



minus end<sup>4</sup>. Thus, by comparing the growth rate at the two microtubule ends with and without added  $\gamma$ TuRC, we should be able to determine whether this complex can prevent the elongation of one end selectively<sup>21</sup>.

To determine growth rates at the two ends, we allowed microtubules to polymerize in two steps. By adding a small amount of rhodamine-labelled tubulin to an excess of purified tubulin, we produced dimly labelled microtubules in both the presence and absence of the purified  $\gamma$ TuRC. A higher concentration of rhodamine-labelled tubulin was then pulsed into the polymerization reaction, resulting in the addition of brightly labelled microtubule segments onto the ends of the dimly labelled microtubules.

We observed three classes of microtubules: (1) uniformly brightly labelled microtubules (Fig. 5e, arrow); (2) microtubules labelled dimly at one end and brightly at the other (Fig. 5f, arrowhead); (3) microtubules labelled dimly in the middle and brightly at both ends (Fig. 5e, arrowhead). As expected, there were more microtubules in each class in the reactions containing  $\gamma$ TuRC. The first class represents microtubules that were nucleated and grew entirely in the second polymerization step; these were not studied further. The second and third classes represent microtubules that elongated from one or both ends of the dimly labelled microtubules, respectively, and these were analysed as follows.

The lengths of the bright microtubule segments on the ends of dim microtubules were measured and plotted. The longer of the two ends was designated as the plus end, which is known to be correct for more than 90% of our control microtubules<sup>22</sup>; the shorter end or the end missing a bright segment was designated as the minus end. Figure 5a and b compares plus-end growth

without and with  $\gamma$ TuRC, respectively. As there is no significant difference between these two sets of data, we can conclude that the  $\gamma$ TuRC has no effect on the growth of the plus ends of microtubules. (These data also confirm that designating the longer end as the plus end in the presence of the  $\gamma$ TuRC did not bias our results.)

Figure 5c and d compares minus-end growth without and with  $\gamma$ TuRC, respectively. We see that, in the presence of  $\gamma$ TuRC, 33.7% (114 of 338) of the total microtubules exhibited no minus-end growth, versus only 4% (4 of 100) in the absence of  $\gamma$ TuRC. Furthermore, in those cases where minus ends were extended, their average lengths ( $3.8 \pm 2.0 \mu\text{m}$ ) were significantly shorter ( $P < 0.001$ ) than the control minus ends ( $5 \pm 1.8 \mu\text{m}$ ). Taken together, these results strongly suggest that the  $\gamma$ TuRC totally blocks the growth of many minus ends, while partially or temporarily inhibiting the elongation of a large fraction of the remaining minus ends of the microtubules that it nucleates.

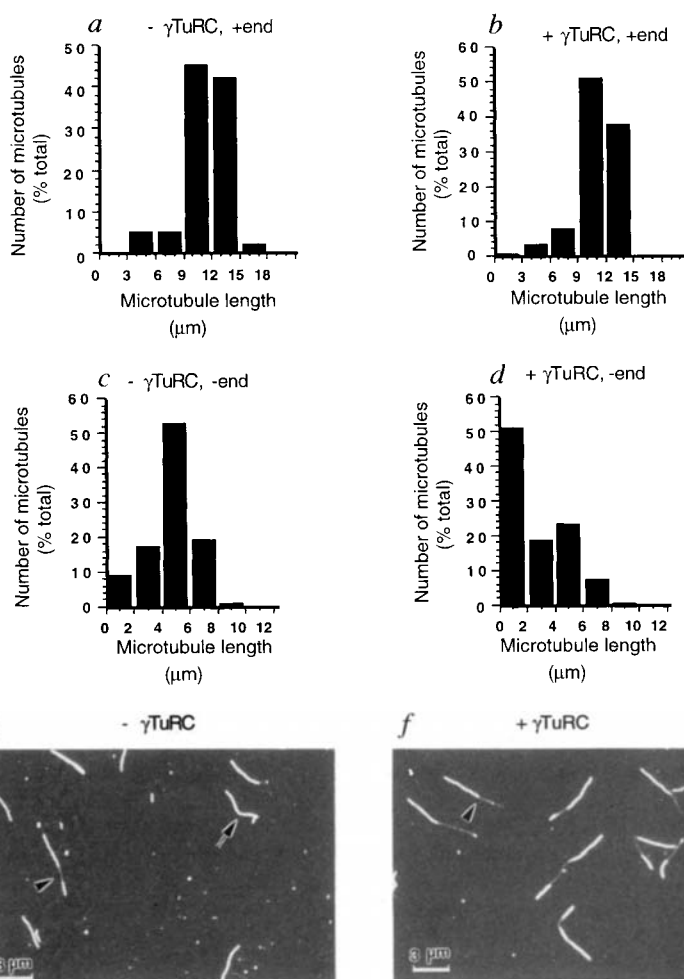
## Discussion

The centrosome has been studied for more than a century<sup>23</sup>, but the mechanism by which it carries out its main function, the nucleation of microtubule growth, has remained a mystery. The discovery of  $\gamma$ -tubulin provided a new handle on this problem<sup>7</sup>. We have now purified a large  $\gamma$ -tubulin-containing protein complex from *Xenopus* in order to study its function.

Like the centrosome, the  $\gamma$ TuRC binds and caps the minus ends of microtubules. In the case of interphase centrosomes, this capping is generally thought to be stable and complete in many cell types; however, for mitotic centrosomes, capping is temporary<sup>24-27</sup>. Although we do not understand what causes this difference between the interphase and mitotic centrosomes,

FIG. 5  $\gamma$ TuRC blocks elongation of the minus end of a microtubule. a–d, Length histograms. a, The length of plus ends produced without  $\gamma$ TuRC (100 microtubules were measured; mean  $11.2 \pm 2.2 \mu\text{m}$ ); b, length of plus ends produced with  $\gamma$ TuRC (338 microtubules were measured; mean  $11.2 \pm 1.9 \mu\text{m}$ ); c, length of minus ends produced without  $\gamma$ TuRC (100 microtubules were measured; mean  $4.8 \pm 2.0 \mu\text{m}$ , including zero length class); d, length of minus ends produced with  $\gamma$ TuRC (338 microtubules were measured; mean  $2.3 \pm 2.4 \mu\text{m}$ , including zero length class). e, A typical field of labelled microtubules produced without  $\gamma$ TuRC; f, typical field of labelled microtubules produced with  $\gamma$ TuRC. Note that the presence of  $\gamma$ TuRC affects the lengths of the minus ends, but not the plus ends of microtubules.

**METHODS.** Faintly fluorescent microtubules were first produced by 1 min incubation at  $37^\circ\text{C}$  of  $25 \mu\text{M}$  tubulin (rhodamine-labelled tubulin and unlabelled tubulin at a ratio of 1 to 30) either in the presence of  $6.4 \mu\text{g ml}^{-1}$   $\gamma$ TuRC (final concentration) or buffer. Then, rhodamine-labelled tubulin was added in enough BRB80 buffer to drop the tubulin concentration to  $11 \mu\text{M}$  and to increase the ratio of the labelled to unlabelled tubulin from 1 to 3. After 4 min at  $37^\circ\text{C}$ , reactions were fixed with 1% glutaraldehyde and diluted into BRB80 containing 60% glycerol. To eliminate background fluorescence from unpolymerized tubulin, microtubules were spun through a glycerol cushion (25% v/v glycerol in BRB80) onto a round coverslip at the bottom of a modified 15 ml Corex tube<sup>6</sup> and mounted with antifade (18 mM Tris-Cl, pH 8.2, with 80% v/v glycerol, 9.1% p-diaminobenzene as free base). Random fields of microtubules were photographed at a final magnification of  $\times 600$  and the lengths of the bright microtubule segments were measured from the negatives.



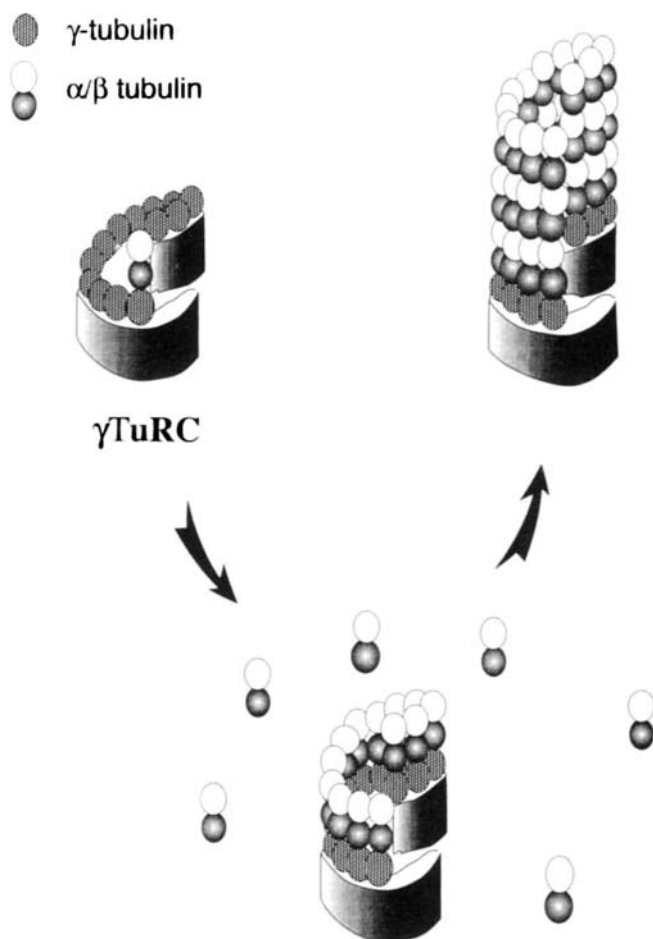


FIG. 6 Seeded nucleation model for microtubule growth. The  $\alpha/\beta$  tubulin dimers bind to the  $\gamma$ TuRC and are stabilized by interaction with  $\gamma$ -tubulins. These assembled tubulin molecules act as a seed to nucleate further microtubule assembly, which determines the protofilament number.

it is interesting that the  $\gamma$ TuRC we purified from a mitotic extract (*Xenopus* eggs are arrested in a metaphase stage of meiosis) causes a partial capping of the minus ends of microtubules (Fig. 5). The association of purified  $\gamma$ -TuRC with the minus ends of microtubules may be reversible, allowing a fraction of the  $\gamma$ TuRC to dissociate from microtubules during our experiments.

Is the  $\gamma$ TuRC the long-sought-after microtubule-nucleating factor in the PCM? Our studies show that the  $\gamma$ TuRC starts to nucleate more microtubules than does a buffer control at 15  $\mu$ M

tubulin (Fig. 3b). To assess the similarity between microtubule nucleation induced by the  $\gamma$ TuRC and that of the centrosome, we have done parallel microtubule-nucleation experiments in the presence of purified  $\gamma$ TuRC or purified *Drosophila* centrosomes. We found that *Drosophila* centrosomes start to nucleate microtubules at a tubulin concentration of 10–15  $\mu$ M (data not shown), which is comparable to the concentration that is effective for purified  $\gamma$ TuRC. Although this is an indirect comparison (between *Xenopus*  $\gamma$ TuRC and the *Drosophila* centrosome), it suggests that the microtubule-nucleating capacity of  $\gamma$ TuRC may be similar to that of centrosomes. As  $\gamma$ -tubulin is found in the PCM and is required for microtubule nucleation from the centrosome<sup>8</sup>, we propose that  $\gamma$ TuRC is the main microtubule-nucleating material in the PCM. This does not exclude the possibility that additional factors at the centrosome may join the  $\gamma$ TuRC and contribute to the microtubule-nucleating ability of the centrosome.

Our proposal is supported by electron microscopy studies<sup>28</sup>, in which a large number of ring structures were observed throughout the PCM of isolated *Drosophila* centrosomes; these rings have a diameter similar to that of the  $\gamma$ TuRC<sup>28</sup>. Furthermore, the rings bind a gold-labelled anti- $\gamma$ -tubulin antibody, and they disappeared when the centrosomes were mixed with tubulin to nucleate the growth of microtubules. After microtubule growth, staining with gold-labelled antibody reveals that  $\gamma$ -tubulin is located at the minus end of each microtubule<sup>29</sup>, which is to be expected if  $\gamma$ TuRC is the microtubule-nucleating factor in the PCM.

The structure of the  $\gamma$ TuRC provides a clue as to how it may nucleate a microtubule. First, the outer diameter of a  $\gamma$ TuRC (25–28 nm) is similar to the outer diameter of a microtubule (25 nm). Second, a  $\gamma$ TuRC has less than two but more than one helical turn, allowing a fit to the end of a microtubule that has a three-start helix with a B-surface lattice<sup>30,31</sup>. Finally, we estimate that there are 1–2  $\alpha/\beta$  tubulin dimers and 10–13  $\gamma$ -tubulin monomers for every  $\gamma$ TuRC. Based on these data, we propose a model that extends that of Oakley *et al.*<sup>9</sup>, which we call the seeded-nucleation model. As depicted in Fig. 6, the non-tubulin proteins (p195, p133, p109 and p75s) in  $\gamma$ TuRC may define the framework of a helical structure to which the  $\gamma$ -tubulins bind. The  $\alpha/\beta$  tubulins and the  $\gamma$ -tubulins then provide a seed for the assembly of more  $\alpha/\beta$  tubulins, which are stabilized by their interactions with  $\gamma$ -tubulin and with each other. Once a short microtubule seed is nucleated, the  $\gamma$ TuRC stabilizes it by capping its minus end.

The model in Fig. 6 predicts that the number of  $\gamma$ -tubulins in the  $\gamma$ TuRC defines the protofilament number of a nucleated microtubule, which is in turn defined by one or more of the larger proteins in the  $\gamma$ TuRC. As specialized tissues exist in which the number of protofilaments in each microtubule is greater than the standard 13 (ref. 32), this feature of the model should be readily testable. □

Received 10 August; accepted 9 November 1995.

- Weisenberg, R. C. *Science* **177**, 1104–1105 (1972).
- Bryan, J. & Wilson, L. *Proc. natn. Acad. Sci. U.S.A.* **68**, 1762–1766 (1971).
- Maldelkow, E., Song, Y. H. & Mandelkow, E. M. *Trends Cell Biol.* **5**, 262–266 (1995).
- Bergen, L. G., Kuriyama, R. & Borisy, G. G. *J. Cell Biol.* **84**, 151–159 (1980).
- Kellogg, D. R., Moritz, M. & Alberts, B. M. *Rev. Biochem.* **63**, 639–674 (1994).
- Evans, L., Mitchison, T. & Kirschner, M. J. *Cell Biol.* **100**, 1185–1191 (1985).
- Oakley, C. E. & Oakley, B. R. *Nature* **338**, 662–664 (1989).
- Oakley, B. R. *Trends in Cell Biol.* **2**, 1–5 (1992).
- Oakley, B. R., Oakley, C. E., Yoon, Y. & Jung, M. K. *Cell* **61**, 1289–1301 (1990).
- Zheng, Y., Jung, M. K. & Oakley, B. R. *Cell* **65**, 817–823 (1991).
- Stearns, T., Evans, L. & Kirschner, M. W. *Cell* **65**, 825–836 (1991).
- Horio, T. *et al. J. Cell Sci.* **99**, 693–700 (1991).
- Joshi, H. C., Palacios, M. J., McNamara, L. & Cleveland, D. W. *Nature* **356**, 80–83 (1992).
- Felix, M.-A., Antony, C., Wright, M. & Maro, B. J. *Cell Biol.* **124**, 19–31 (1994).
- Mitchison, T. J. *Science* **261**, 1044–1047 (1993).
- Vassilev, A. *et al. J. Cell Sci.* **108**, 1083–1092 (1995).
- Melki, R., Vainberg, I. E., Chow, R. L. & Cowan, N. J. *J. Cell Biol.* **122**, 1301–1310 (1993).
- Raff, J. W., Kellogg, D. R. & Alberts, B. M. *J. Cell Biol.* **121**, 823–835 (1993).
- Stearns, T. & Kirschner, M. W. *Cell* **76**, 623–637 (1994).
- Siegel, L. M. & Monty, K. J. *Biochim. biophys. Acta* **112**, 346–362 (1966).
- Hyman, A. A. & Mitchison, T. J. *Nature* **351**, 206–211 (1991).
- Mitchison, T. J. & Kirschner, M. W. *J. Cell Biol.* **101**, 755–765 (1985).

- Wilson, E. B. in *The Cell in Development and Heredity* (ed. Moore, J. A.) 1232 (Macmillan, New York, 1925).
- Belmont, L. D., Hyman, A. A., Sawin, K. E. & Mitchison, T. J. *Cell* **62**, 579–589 (1990).
- Ault, J. G. & Nicklas, R. B. *Chromosoma* **98**, 33–39 (1989).
- Kitanishi-Yumura, T. & Fukui, Y. *Cell. Motil. Cytoskel.* **8**, 106–117 (1987).
- Harris, P., Osborn, M. & Weber, K. J. *Cell Biol.* **84**, 668–678 (1980).
- Moritz, M. *et al. J. Cell Biol.* **130**, 1149–1159 (1995).
- Moritz, M., Braunfeld, M. B., Sedat, J. W., Alberts, B. M. & Agard, D. A. *Nature* **378**, 638–640.
- Song, Y. H. & Mandelkow, E. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1671–1675 (1993).
- Kikkawa, M., Ishikawa, T., Nakata, T., Wakabayashi, T. & Hirokawa, N. *J. Cell Biol.* **127**, 1965–1971 (1994).
- Tucker, J. B. *et al. Eur. J. Cell Biol.* **41**, 279–289 (1986).
- Harlow, E. & Lane, D. in *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1988).
- Murray, A. W., Solomon, M. J. & Kirschner, M. W. *Nature* **339**, 280–286 (1989).
- Hyman, A. A. *et al. Meth. Enzym.* **196**, 478–487 (1991).
- Gill, S. C. & von Hippel, P. H. *Analyt. Biochem.* **182**, 319–326 (1989).

ACKNOWLEDGEMENTS. We thank J. Barry for advice on protein purification, T. Hirano for advice on preparing the *Xenopus* egg extract, M. Moritz, K. Oogema, B. Keilum and J. Barry for comments on the manuscript, R. Kline for Fig. 6, and M. Moritz for *Drosophila* centrosomes. This work was supported by separate grants from the NIH to B.A. and T.M., and by the Howard Hughes Medical Institute (M.L.W.).

# Structure and ligand recognition of the phosphotyrosine binding domain of Shc

Ming-Ming Zhou<sup>\*</sup>, Kodimangalam S. Ravichandran<sup>†</sup>, Edward T. Olejniczak<sup>\*</sup>, Andrew M. Petros<sup>\*</sup>, Robert P. Meadows<sup>\*</sup>, Michael Sattler<sup>\*</sup>, John E. Harlan<sup>\*</sup>, Warren S. Wade<sup>\*</sup>, Steven J. Burakoff<sup>†</sup> & Stephen W. Fesik<sup>\*‡</sup>

<sup>\*</sup> Pharmaceutical Discovery Division, Abbott Laboratories, Abbott Park, Illinois 60064, USA

<sup>†</sup> Division of Pediatric Oncology, Dana-Farber Cancer Institute, and Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115, USA

<sup>‡</sup> To whom correspondence should be addressed

**The nuclear magnetic resonance structure of the phosphotyrosine binding (PTB) domain of Shc complexed to a phosphopeptide reveals an alternative means of recognizing tyrosine-phosphorylated proteins. Unlike in SH2 domains, the phosphopeptide forms an antiparallel  $\beta$ -strand with a  $\beta$ -sheet of the protein, interacts with a hydrophobic pocket through the (pY-5) residue, and adopts a  $\beta$ -turn. The PTB domain is structurally similar to pleckstrin homology domains (a  $\beta$ -sandwich capped by an  $\alpha$ -helix) and binds to acidic phospholipids, suggesting a possible role in membrane localization.**

The adaptor protein Shc is involved in signalling by means of many receptors. Shc has been implicated in Ras-dependent MAP kinase activation following the stimulation of growth factor (insulin and epidermal (EGF), nerve (NGF) and platelet-derived growth factor (PDGF))<sup>1-6</sup>, cytokine (interleukin (IL)-2, IL-3, IL-5, erythropoietin (EPO) and granulocyte macrophage colony-stimulating factor (GM-CSF))<sup>6-8</sup> and antigen (T and B cell)<sup>9,10</sup> receptors. On receptor stimulation, Shc binds to activated, tyrosine-phosphorylated receptors and becomes phosphorylated. Tyrosine-phosphorylated Shc subsequently interacts with the Src homology 2 (SH2) domain of Grb2, which binds by its SH3

domains to the guanine nucleotide exchange factor, SOS, providing the link to the Ras activation pathway<sup>11</sup>.

The role of Shc in signal transduction processes has recently expanded to include pathways that are activated by the stimulation of G-protein-coupled receptors or elevated levels of intracellular calcium<sup>12,13</sup>. These pathways are mediated by the  $\beta\gamma$  subunits of heterotrimeric G proteins ( $G\beta\gamma$ )<sup>12</sup> and stimulated by the calcium-dependent activation of a newly discovered protein tyrosine kinase, PYK2 (ref. 13). As in receptor tyrosine kinase (RTK) signalling, activation of these pathways results in the phosphorylation of Shc, which leads to the formation of the



Fig. 1 **a**, Sequence alignment of putative PTB domains. The sequences were aligned by considering the NMR-derived secondary structure of the PTB domain of the Shc (green,  $\alpha$ -helices; yellow,  $\beta$ -strands) and key homologous residues (red) in the tertiary structure. Residues of the PTB domain of Shc that show NOEs to the TrkA phosphopeptide are indicated by asterisks. **b**, Secondary structure of the PTB domain. NOEs that define the structure of the  $\beta$ -sheet are indicated by arrows. Hydrogen bonds supported by the amide exchange data that were included in the structure calculations are indicated by broken lines. The phosphopeptide residues are shown in red.



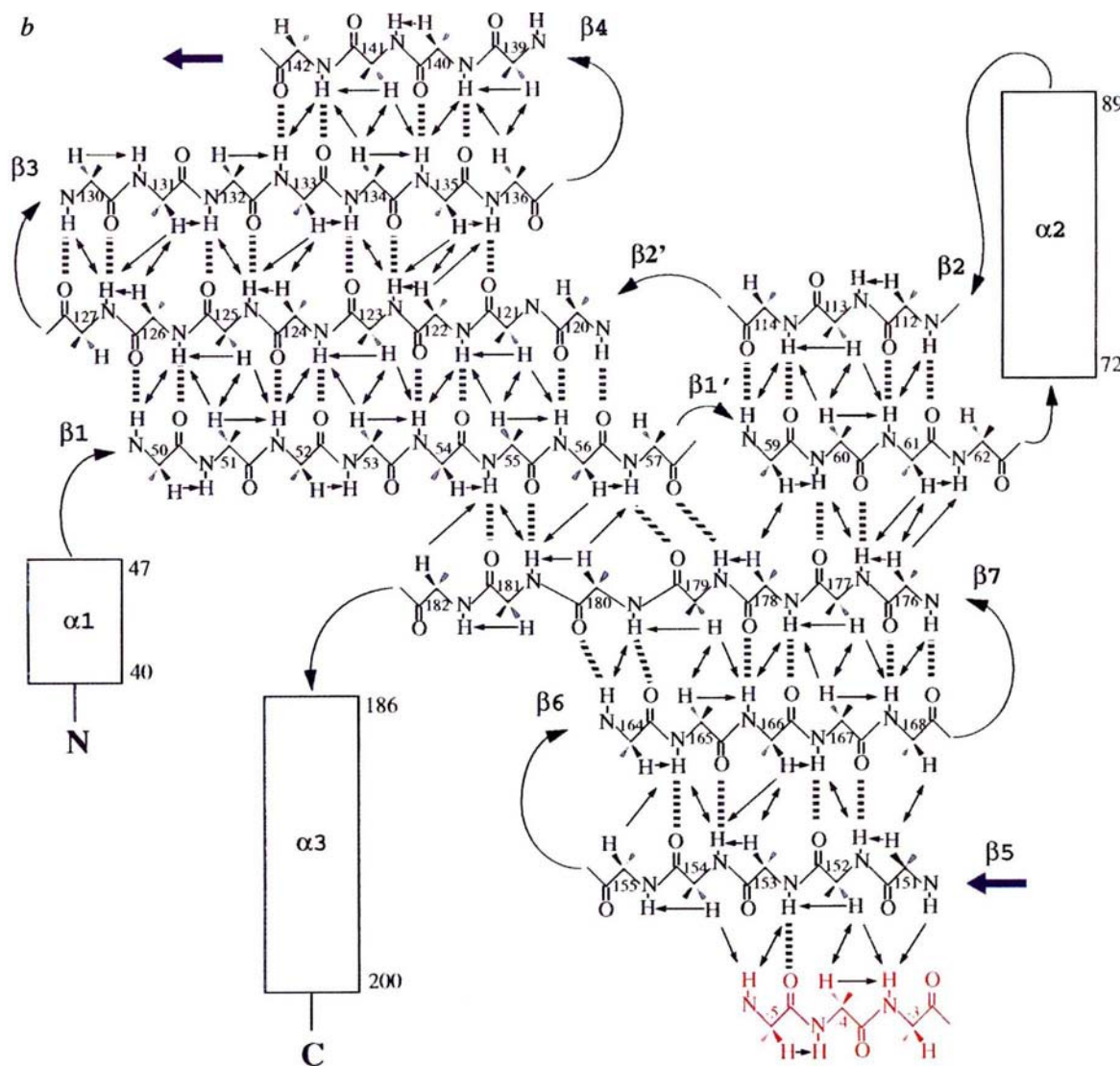
Shc:Grb2:SOS complex and Ras activation<sup>12,13</sup>. However, in contrast to RTK signalling in which Shc binds to membrane-associated tyrosine-phosphorylated receptors, the manner in which Shc is recruited to the membrane in G $\beta$ -mediated signal transduction is unknown.

The Shc protein contains three domains: an amino-terminal region called the phosphotyrosine binding (PTB) domain<sup>14-16</sup>, a collagen homology domain that contains the binding site (Tyr 317) for Grb2, and a carboxy-terminal Src-homology 2 (SH2) domain. Although both the SH2 and PTB domains of Shc can bind to tyrosine-phosphorylated proteins, these two domains exhibit markedly different phosphopeptide-binding specificities<sup>17,22</sup>. Apart from the presence of an F/YLVR sequence, there is no obvious sequence homology between the PTB and SH2 domains. Furthermore, unlike SH2 domains, the PTB domain of Shc recognizes amino acid residues N-terminal (rather than C-terminal) to the phosphotyrosine, and preferentially interacts with  $\Phi$ XNPXPY-containing sequences in which the (pY-5) residue ( $\Phi$ ) is hydrophobic<sup>19</sup>. Indeed, the PTB domain of Shc binds specifically to NPXPY-containing sequences found in EGF, Her2/neu, NGF<sup>15,23</sup>, insulin receptors<sup>16,24</sup>, p145 (ref. 14) and polyomavirus middle T antigen<sup>19</sup>. The PTB domain is also found in several other proteins<sup>14,21,25,26</sup> (Fig. 1a) and may play an important role for their function. However, no three-dimensional structure of the PTB domain has been reported, either alone or when complexed to a tyrosine-phosphorylated peptide.

Here we describe the three-dimensional structure of the PTB domain of Shc complexed to a tyrosine-phosphorylated peptide of 12 residues that is derived from the cytoplasmic juxta-membrane domain of the nerve growth factor receptor TrkA, as determined by heteronuclear multidimensional nuclear magnetic resonance (NMR) spectroscopy. This study provides the first view of the PTB domain's structure and the way in which the NPXPY-containing peptides are recognized by this newly discovered domain. Based on the structure and site-directed mutagenesis, we have also identified key residues that participate in binding to tyrosine-phosphorylated proteins. These data should be useful in defining the role of the PTB domain in biological processes mediated by Shc and the other proteins that also contain this domain. Finally, we show that the PTB domain is structurally similar to PH domains and binds to acidic phospholipids, which may have the functional significance of recruiting Shc to the membrane in G $\beta$ -mediated, Ras-dependent activation of MAP kinase.

### Structure determination

The NMR studies were conducted using a 1:1 complex of an N-terminal fragment of Shc (residues 17-207) and a 12-residue tyrosine-phosphorylated peptide (HIENPQpYFSDA) derived from the sole Shc binding site (phosphotyrosine 490) of the nerve growth factor receptor TrkA. This phosphopeptide has previously been shown to bind tightly ( $K_D = 53$  nM) to the Shc



PTB domain (residues 17–207)<sup>22</sup>. The details of the NMR sample preparation and structure determination are provided (Fig. 2 legend). The first step in the structure determination was to assign the NMR signals. For Shc, the backbone and side-chain assignments of the protein were greatly facilitated by the use of uniformly <sup>15</sup>N-, <sup>13</sup>C-labelled protein that was fractionally (75%) deuterated. Because of the favourable <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N relaxation rates caused by the partial deuterium labelling, constant time spectra with higher digital resolution<sup>27</sup> that were critical for making the resonance assignments of the protein were acquired. In addition, many more intermolecular nuclear Overhauser enhancements (NOEs) used in defining the structure of the complex were detected in <sup>13</sup>C-edited ( $F_1$ ), <sup>13</sup>C-, <sup>15</sup>N-filtered ( $F_3$ ) three-dimensional NOE experiments using the fractionally deuterated versus fully protonated protein, particularly intermolecular NOEs involving  $\beta$ -CH<sub>2</sub> protons, which rapidly relax in proteins of this size (manuscript in preparation).

The three-dimensional structure of the PTB domain of Shc complexed with the TrkA phosphopeptide was determined using a distance geometry/simulated annealing protocol<sup>28,29</sup> from a total of 2,121 NMR-derived restraints, including 128 restraints derived from intermolecular NOEs between the protein and phosphopeptide. As indicated in Table 1, the structures satisfy the distance restraints with no violation greater than 0.4 Å, and exhibit good covalent geometry and non-bonded contacts. Figure 2a depicts a stereo view of the backbone superposition of the 15 lowest energy structures of the PTB domain for residues 35–207 and the TrkA phosphopeptide. For clarity, the N-terminal residues of the PTB domain (17–34) that were ill-defined by the NMR data are not shown. These residues do not exhibit any long-range NOEs to the rest of the protein and are probably not part of the PTB domain. The backbone structure for the remainder of the protein is well defined by the NMR data, with the exception of the last three residues at the C terminus and one of the loops (residues 90–108). Excluding this loop, the atomic r.m.s. distribution about the mean coordinate positions for residues 40–207 of the PTB domain and the phosphopeptide (residues pY–7 to pY+1) is  $0.69 \pm 0.13$  Å for the backbone atoms and  $1.34 \pm 0.16$  Å for all heavy atoms.

## PTB domain structure

Figure 1b depicts the secondary structure of the PTB domain, along with the supporting NOE and amide-exchange data. The secondary structure as determined by NMR is largely consistent with previously reported predictions<sup>25</sup>. The tertiary structure of the PTB domain of Shc consists of a  $\beta$ -sandwich containing two nearly orthogonal, anti-parallel  $\beta$ -sheets and three  $\alpha$ -helices (Fig. 2b). One of the  $\beta$ -sheets comprises the first four strands ( $\beta 1$ – $\beta 4$ ), and the other sheet comprises the last three strands ( $\beta 5$ – $\beta 7$ ) together with part of the first and second strands. The first  $\beta$ -strand is split ( $\beta 1$  and  $\beta 1'$ ), and forms a right-handed  $\beta$ -bend. A highly conserved glycine (Gly 58)<sup>25</sup> (Fig. 1a) is located in the middle of the  $\beta$ -bend and forms a kink that maintains the registration of the hydrogen bonds between the  $\beta$ -strands. The second  $\beta$ -strand is also split ( $\beta 2$  and  $\beta 2'$ ) and forms a bend. In this case, a bulge is formed at the bend, consisting of residues 115–119. The  $\alpha$ -helix ( $\alpha 2$ ) that connects  $\beta 1$  and  $\beta 2$  is located on one side of the second  $\beta$ -sheet, and covers the hydrophobic residues, V60 and I179, that are found on the outside of this sheet. The C-terminal amphipathic  $\alpha$ -helix ( $\alpha 3$ ) (residues 186–200) caps one end of the barrel-like structure with the side chains of V190, I191, I194 and F198 of the helix pointing into the hydrophobic core. The short N-terminal  $\alpha$ -helix ( $\alpha 1$ , residues 40–47) packs against the C-terminal  $\alpha$ -helix at an angle of  $\sim 60^\circ$  and forms predominantly hydrophilic interactions with  $\alpha 3$ .

## Phosphopeptide binding site

The TrkA phosphopeptide binding site is composed of an elongated cleft formed by  $\beta 5$ , the C-terminal helix ( $\alpha 3$ ), and the

TABLE 1 Structural statistics and r.m.s. deviations for 15 NMR-derived structures of the SHC PTB domain

| Structural statistics                                  | <SA>              | <SA> <sub>r</sub> |
|--------------------------------------------------------|-------------------|-------------------|
| r.m.s.d. from experimental distance restraints (Å)*    |                   |                   |
| intra-residue (595)                                    | $0.014 \pm 0.006$ | 0.015             |
| sequential (558)                                       | $0.017 \pm 0.002$ | 0.022             |
| medium range (189)                                     | $0.022 \pm 0.005$ | 0.019             |
| long range (450)                                       | $0.018 \pm 0.002$ | 0.016             |
| intermolecular (128)                                   | $0.012 \pm 0.004$ | 0.011             |
| hydrogen bonds (150)                                   | $0.036 \pm 0.003$ | 0.035             |
| r.m.s.d. from experimental torsional restraints (deg)† |                   |                   |
| $\Phi$ angles (51)                                     | $0.49 \pm 0.15$   | 0.41              |
| X-PLOR potential energy (kcal mol <sup>-1</sup> )      |                   |                   |
| $E_{\text{tot}}$                                       | $369 \pm 30$      | 386               |
| $E_{\text{bond}}$                                      | $28 \pm 1$        | 29                |
| $E_{\text{ang}}$                                       | $222 \pm 14$      | 235               |
| $E_{\text{imp}}$                                       | $25 \pm 2$        | 25                |
| $E_{\text{repel}}$                                     | $51 \pm 7$        | 57                |
| $E_{\text{noe}}$                                       | $38 \pm 8$        | 39                |
| $E_{\text{cdih}}$                                      | $0.8 \pm 0.4$     | 0.5               |
| $E_{\text{LJ}}^\ddagger$                               | $-940 \pm 14$     | -931              |
| r.m.s.d. from ideal values                             |                   |                   |
| bonds (Å)                                              | $0.00 \pm 0.00$   | 0.00              |
| angles (deg)                                           | $0.50 \pm 0.02$   | 0.51              |
| impropers (deg)                                        | $0.45 \pm 0.02$   | 0.48              |
| cartesian coordinate r.m.s.d. (Å)                      |                   |                   |
| <SA> versus <SA> <sub>r</sub> <sup>§</sup>             | $1.01 \pm 0.13$   | $1.75 \pm 0.13$   |
| <SA> versus <SA> <sub>r</sub> <sup>  </sup>            | $0.69 \pm 0.13$   | $1.34 \pm 0.16$   |
| <SA> versus <SA> <sub>r</sub> <sup>¶</sup>             | $0.56 \pm 0.15$   | $1.03 \pm 0.36$   |
| <SA> versus <SA> <sub>r</sub> <sup>#</sup>             | $0.27 \pm 0.15$   | $0.92 \pm 0.38$   |

<SA> is the ensemble of 15 NMR-derived solution structures; <SA><sub>r</sub> is the mean atomic structure obtained by averaging the coordinates of the individual <SA> structures following a least-squares superposition of the backbone heavy atoms for residues 40–207 and the phosphopeptide (pY–7 to pY+1); and <SA><sub>r</sub> is the energy-minimized average structure; 1 kcal = 4.18 kJ. The X-PLOR  $F_{\text{repel}}$  function<sup>28</sup> was used to simulate van der Waals interactions with a force constant of 4.0 kcal mol<sup>-1</sup> Å<sup>-4</sup> with the atomic radii set to 0.80 times their CHARMM<sup>42</sup> values.

\* A total of 1920 non-trivial NOE-derived distance restraints were used with a square-well potential ( $F_{\text{noe}} = 50$  kcal mol<sup>-1</sup> Å<sup>-2</sup>). In addition, 75 hydrogen bonds were included and given bounds of 1.8–2.4 Å (H→O), and 2.8–3.2 Å (N→O). No distance restraint was violated by more than 0.4 Å in any of the final structures. Medium-range NOEs are observed between protons separated by more than one and less than five residues in the primary sequence. Long-range NOEs are observed between protons separated by five or more residues.

† Torsional restraints were applied to 51  $\Phi$  angles with values of  $-120 \pm 40^\circ$  for those angles with <sup>3</sup>J<sub>H<sub>N</sub>H<sub>H</sub></sub> coupling constants >9.0 Hz and  $-60 \pm 40^\circ$  for coupling constants <5.5 Hz. Restraints for the latter were applied only in  $\alpha$ -helical regions. Force constants of 200 kcal mol<sup>-1</sup> rad<sup>-2</sup> were applied for all torsional restraints.

‡ The Lennard-Jones potential was not used during any refinement stage.

§ r.m.s.d. for residues 40–207 and the phosphopeptide (pY–7 to pY+1).

|| r.m.s.d. for residues 40–90, 110–207 and the phosphopeptide (pY–7 to pY+1).

¶ Atomic r.m.s.d. of the phosphopeptide when only the protein backbone atoms (residues 40–207) were least-squares fitted.

# r.m.s.d. of the phosphopeptide (pY–7 to pY+1) when superimposed upon the average coordinates for these residues.

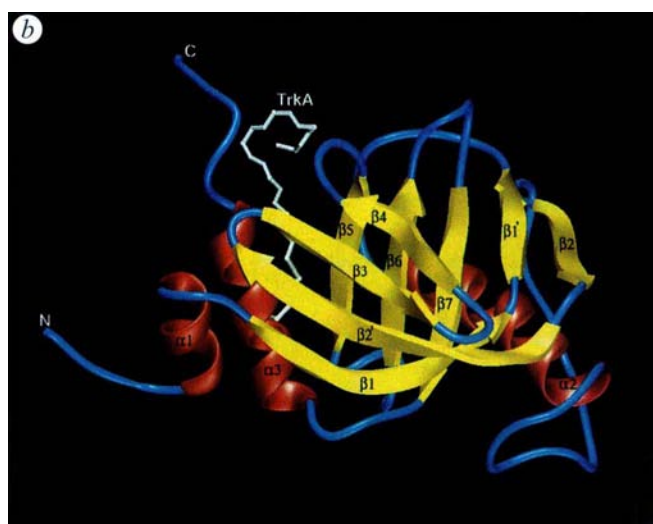
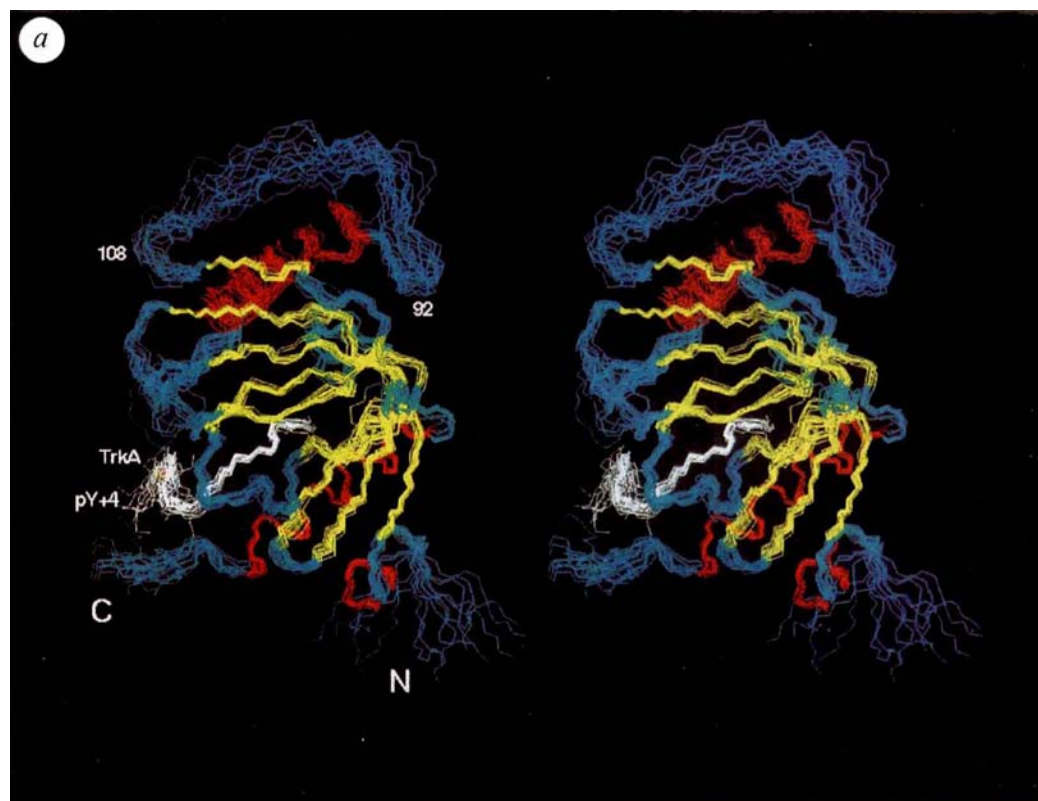
loop connecting  $\beta 1$  and  $\alpha 2$  (Fig. 2b). The N terminus of the phosphopeptide adopts an extended conformation and forms an additional antiparallel  $\beta$ -strand of the  $\beta$ -sheet (Figs 1b and 3a). The side chain of Ile (pY–5) shows numerous contacts with the protein (41 intermolecular NOEs), and binds in a hydrophobic pocket between  $\beta 5$  and  $\alpha 3$  that is composed of Ile 191, Ile 194, and Phe 198 (Fig. 3a, b). This explains the preference for a large

FIG. 2 *a*, Stereo view of the backbone (N, Ca, C') of 15 superimposed NMR-derived structures of the PTB domain of Shc (residues 35–207) complexed to the TrkA phosphopeptide. *b*, Ribbons<sup>13</sup> depiction of the averaged minimized NMR structure of the PTB domain/phosphopeptide complex. The  $\beta$ -strands are shown in yellow, the  $\alpha$ -helices are in red, and the phosphopeptide is in white. METHODS. The N-terminal PTB domain of Shc (residues 17–207) was cloned, expressed and purified as described previously<sup>22</sup>. The protein was subcloned into the pET15b expression vector and expressed in *Escherichia coli* BL21(DE3) cells which were co-transformed with the tRNA plasmid Pom61(tetR). Uniformly <sup>15</sup>N- and <sup>15</sup>N-, <sup>13</sup>C-labelled proteins were prepared for the NMR experiments by growing bacteria that overexpress the PTB domain in a minimal medium containing <sup>15</sup>NH<sub>4</sub>Cl with or without [U-<sup>13</sup>C]glucose. A uniformly <sup>15</sup>N-, <sup>13</sup>C-labelled and frac-

tionally deuterated protein sample was also prepared by growing the cells in 75% <sup>2</sup>H<sub>2</sub>O. The Shc PTB domain was purified by affinity chromatography on a nickel-IDA column (Invitrogen) followed by the removal of the His tag by thrombin cleavage. The final purification of the protein was achieved by cation-exchange chromatography. The tyrosine-phosphorylated peptide of TrkA was prepared on an ABI (Applied Biosystems Inc.) 430A peptide synthesizer using Fmoc/HBTU chemistry. Phosphotyrosine was incorporated using the reagent Fmoc-Tyr(PO<sub>3</sub>H<sub>2</sub>) with HBTU/HOAt activation. NMR samples of the complex were formed by mixing the labelled protein and the unlabelled phosphopeptide purified by high-performance liquid chromatography (HPLC) in an H<sub>2</sub>O/<sup>2</sup>H<sub>2</sub>O (9/1) or <sup>2</sup>H<sub>2</sub>O Tris buffered solution (20 mM, pH 6.5) containing 20-mM perdeuterated DTT. All NMR spectra were acquired at 35 °C on a Bruker DMX 500 or AMX600 NMR spectrometer. The <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N resonances of the backbone were assigned using the uniformly <sup>15</sup>N-, <sup>13</sup>C-labelled and fractionally deuterated protein by correlating the amide <sup>1</sup>H and <sup>15</sup>N resonances of each amino acid to the Ca, C $\beta$  and C' signals of the *i* and (*i*-1) residues from a series of deuterium-decoupled three-dimensional NMR experiments (HNCA, HN(CO)CA, HN(CA)CB, HN(COCA)CB, HNCO, HN(CA)CO)<sup>27</sup>. The side-chain signals were assigned from 3D and 4D HCCH total correlation spectroscopy (TOCSY), HC(CO)NH-TOCSY, and HBHA(CO)NH experiments<sup>44</sup>. Structures of the PTB domain complexed with the TrkA phosphopeptide were generated with a distance geometry/simulated annealing (DG/SA) protocol<sup>28,29</sup> using the X-PLOR program<sup>45</sup>. The structure calculations used 1,920 proton/proton distance restraints obtained from <sup>15</sup>N- and <sup>13</sup>C-resolved and filtered three-dimensional NOE spectra, 75 hydrogen bonds from an analysis of the amide exchange rates measured from a series of <sup>1</sup>H/<sup>15</sup>N HSQC spectra recorded after the addition of <sup>2</sup>H<sub>2</sub>O, and 51  $\Phi$  angle restraints from the <sup>3</sup>J<sub>H<sub>N</sub>H<sub>α</sub></sub> coupling constants measured in a

hydrophobic residue at the (pY-5) position for tight binding to the PTB domain<sup>19</sup>. Ile (pY-6) also interacts with a hydrophobic region of the protein consisting of Leu 69 and Phe 152. However, the side chains of Ile (pY-6) and His (pY-7) are more solvent exposed than Ile (pY-5), which may explain the lack of binding specificity observed for the residues at these positions<sup>19</sup>.

The NPQpY residues of the TrkA phosphopeptide form a  $\beta$ -turn (Fig. 3*a, b*) consistent with the relatively large NOE observed between the amide protons of pTyr and Gln (pY-1).



HNHA-J experiment<sup>46</sup>. The NOE-derived restraints were categorized as strong (3 Å), medium (4 Å) or weak (5 Å) based on the observed NOE intensities.

A  $\beta$ -turn within the phosphopeptide may be required for the simultaneous interaction of the PTB domain with both the N-terminal residues of the peptide and the phosphotyrosine residue. The formation of this  $\beta$ -turn may also explain, in part, the preference for an Asn and Pro at the (pY-3) and (pY-2) positions, as these residues are highly favoured at the *i* and *i*+1 positions, respectively, of  $\beta$ -turns in protein<sup>30</sup>. Even small peptides containing an NPXY or NPXpY sequence have been found to adopt a stable  $\beta$ -turn in solution<sup>19,31</sup>, including the TrkA phosphopeptide



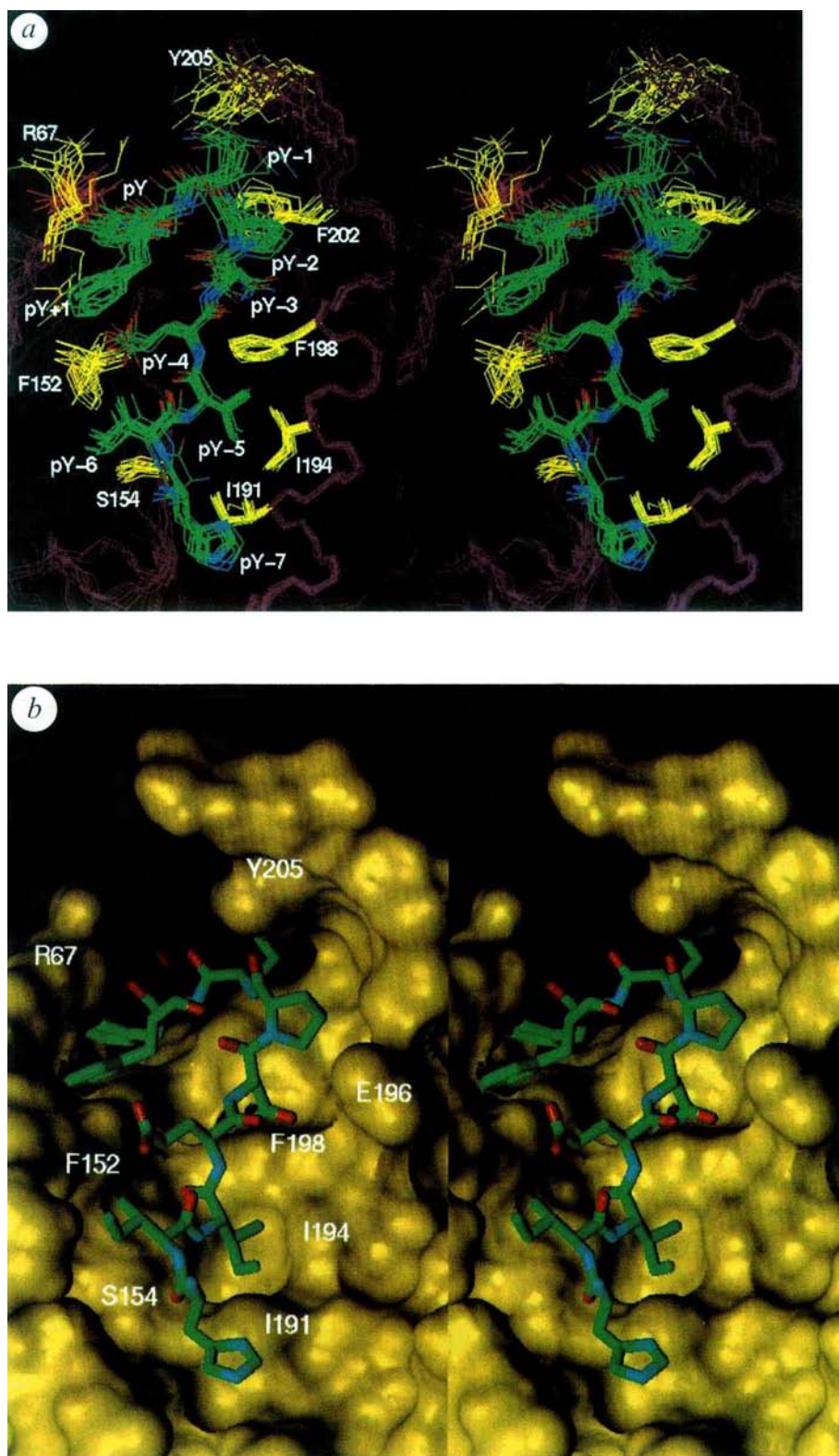
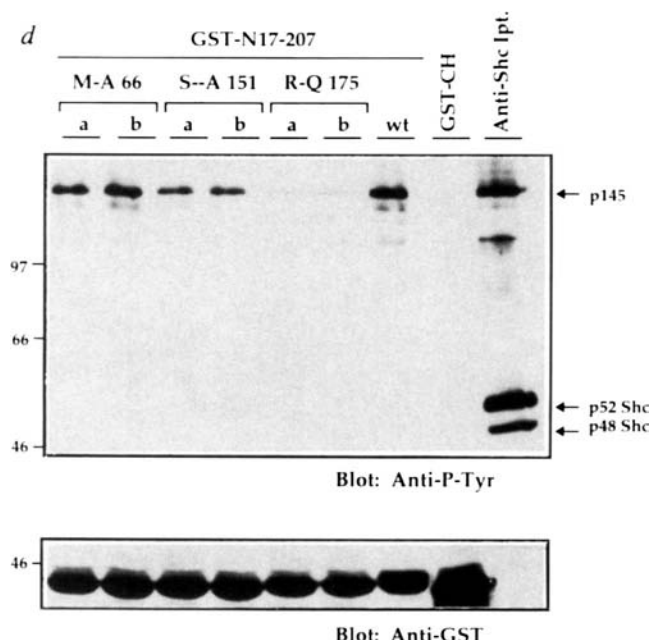
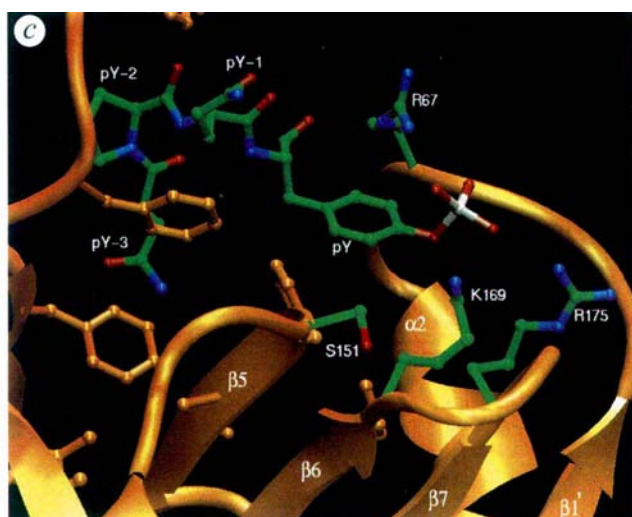


FIG. 3 *a*, Stereo view of the peptide binding site of 15 superimposed NMR-derived structures of the PTB domain/TrkA phosphopeptide complex. The TrkA phosphopeptide (residues pY + 1 to pY - 7) is colour-coded by atom type, and the backbone and selected side chains of the Shc PTB domain are shown in purple and yellow, respectively. *b*, Solid-rendering stereo-view of the phosphopeptide binding pocket (yellow) with the bound TrkA phosphopeptide (colour-coded by atom type). *c*, Ribbons<sup>43</sup> depiction of a portion of the PTB domain of Shc complexed with the TrkA phosphopeptide (residues NPQpY). The phosphopeptide and selected side chains of the protein are colour-coded by atom type. *d*, Binding of wild-type and mutant GST-Shc (17–207) proteins to p145, assessed by anti-phosphotyrosine immunoblotting. Two independent clones (*a* and *b*) were analysed for each mutant.

**METHODS.** Site-directed mutagenesis of Met 66 to Ala, Ser 151 to Ala, Arg 175 to Gln and Lys, and Arg 67 to Gln was performed by polymerase chain reaction (PCR) with primers carrying the desired mutation and cloned into pGex2T (Pharmacia). The presence of appropriate mutations was confirmed by DNA sequencing. The proteins were expressed in *E. coli* and purified using glutathione agarose beads. Fusion proteins (2  $\mu$ g) were incubated with K562 lysates ( $20 \times 10^6$  cell equivalents), and the bound proteins were resolved by 8% sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and analysed by anti-phosphotyrosine immunoblotting<sup>22</sup>. Rough quantitation of the bands was performed by densitometry, and the percentage loss in p145 binding was estimated compared to the wild-type protein. Anti-GST antibody (SantaCruz Biotechnology) was used in immunoblotting, and anti-Shc immunoprecipitation was performed as described previously<sup>22</sup>.

(data not shown). These results suggest that the conformation of the peptide is largely preformed before interacting with the PTB domain. Although it has been suggested that  $\beta$ -turn formation is stabilized by an interaction between the  $\text{NH}_2$  group of the Asn and the aromatic ring of the  $i+3$  residue of the turn<sup>31</sup>, these groups are too far apart to form an interaction in the bound phosphopeptide (Fig. 3*a, b*). Instead, the  $\text{NH}_2$  of the Asn (pY-3) side chain is in close proximity to the aromatic ring of

Phe 198 (Fig. 3*a, b*), which is consistent with the large changes in chemical shifts observed for these protons (7.49/6.80 to 8.41/6.26 p.p.m.) when it binds to the PTB domain. The Asn  $\text{NH}_2$  of the phosphopeptide is also near the carbonyl oxygen of Phe 198. Thus, as well as stabilizing  $\beta$ -turn formation, the Asn at the (pY-3) position may be favoured as a result of an interaction of its side-chain amide with both the aromatic side chain and the carbonyl of Phe 198.



The phosphotyrosine of the peptide interacts with a positively charged site on the protein composed of Arg 67, Ser 151, Lys 169 and Arg 175 (Fig. 3e). To probe the relative contributions of these residues for binding to tyrosine-phosphorylated proteins, we generated PTB domains carrying point mutations of Met 66, Arg 67, Ser 151 and Arg 175. These proteins were expressed as glutathione *S*-transferase (GST) fusion proteins, and their ability to bind tyrosine-phosphorylated p145 from K562 cell lysates was assessed (Fig. 3d). Previously, p145 has been shown to bind specifically to the PTB domain of Shc after a variety of stimuli<sup>14</sup>. Compared with the wild-type protein, mutation of Met 66 and Ser 151 to Ala resulted in a loss of p145 binding by 12% and 41%, respectively (Fig. 3d). Mutation of Arg 67 to Gln also had a partial effect on PTB binding to p145 (36% loss, data not shown). However, mutation of Arg 175 to Gln (Fig. 3d) or Lys (data not shown) completely eliminated binding to p145, ascribing a crucial role for Arg 175 in the interaction with the phosphotyrosine. These results are consistent with the structural information obtained with the PTB-phosphopeptide complex. The phosphotyrosine of the TrkA peptide is in close proximity

to Ser 151 and Arg 175, and the Phe (pY+1) residue of the TrkA peptide forms a hydrophobic interaction with the aliphatic side chain of Arg 67 (Fig. 3c). Although small NOEs were observed between the Met 66 methyl group and the pTyr and Phe residues of the peptide, Met 66 does not contribute significantly to the binding of the phosphopeptide, based on the mutagenesis results (Fig. 3d).

The other residues (SDA) of the phosphopeptide that are C-terminal to the phosphotyrosine do not display any intermolecular NOEs with the PTB domain, are solvent-exposed, and are ill-defined by the NMR data. This is consistent with the observation that binding affinities are not affected by deletion of these amino acid residues<sup>19</sup>.

A previous database search and sequence alignment of putative PTB domains indicated that Ser 151 is invariant in all of the proteins retrieved from the search<sup>25</sup>. Based on the structure of the complex and results from site-directed mutagenesis, it is expected that Arg 67, Lys 169 and Arg 175 would also be conserved in the PTB domains of other proteins. Indeed, in the sequence alignment of the proteins shown in Fig. 1a, all of these residues that participate in the binding of the phosphopeptide are conserved. However, in other putative PTB domains identified from a database search<sup>25</sup>, homologous residues could not be readily found in similar positions after several attempts at realigning the sequences (data not shown). These results suggest either that residues from other parts of the protein substitute for these amino acids, or that the three-dimensional structures of these proteins are significantly different from the PTB domain of Shc.

### Comparison with SH2 domains

As with SH2 domains<sup>32-34</sup>, the phosphopeptide binding site in the PTB domain consists of an antiparallel  $\beta$ -sheet flanked by two  $\alpha$ -helices (Fig. 4a, b). In addition, the phosphotyrosine binding pocket is composed of the same residue types in both domains. The phosphotyrosine binding pocket consists of Arg 67, Arg 175, Lys 169 and Ser 151 in the PTB domain (Fig. 3c), which closely corresponds to Arg  $\alpha$ A2, Arg  $\beta$ B5, Lys  $\beta$ D6 and Ser BC2 in the lck SH2 domain<sup>32</sup>. Moreover, the Arg residue that is required for binding to the phosphotyrosine in both domains (Arg 175 in the PTB domain and Arg  $\beta$ B5 in the SH2 domain) projects from a similar position (Fig. 4a, b).

In general, however, there are more differences than similarities between the SH2 and PTB domains. Although both domains possess a hydrophobic site for binding to the phosphopeptide, the PTB domain forms a hydrophobic interaction with the pY-5 residue while SH2 domains bind to pY+3. Furthermore, in SH2 domains, the phosphopeptide binds almost perpendicular to the  $\beta$ -sheet, whereas it binds parallel to the  $\beta$ -sheet in the PTB domain (Fig. 4a, b). In addition, unlike in SH2 domains, the phosphopeptide adopts a  $\beta$ -turn when bound to the PTB domain. Other differences include the relative orientation of the  $\beta$ -strands and the location of the  $\alpha$ -helices relative to the second  $\beta$ -sheet.

One of the similarities of the two domains that has been cited<sup>14,21</sup> is the presence of a conserved F/YLVR sequence containing Arg  $\beta$ B5. The presence of this motif in the PTB domain has led to the hypothesis that the two domains may be similar for at least this part of the molecule<sup>21</sup>. However, in the PTB domain, the YLVR sequence is far removed from the phosphopeptide binding site (Fig. 4a). This explains why the mutation of this Arg to a Leu does not affect the binding of the PTB domain to p145 (ref. 14).

### Comparison with PH domains

The overall topology of the PTB domain is found in several protein families such as the lipocalcins, fatty-acid binding proteins, and streptavidins<sup>35</sup>, but it is most similar to the PH domains (Fig. 5a). The structures of PTB and PH domains<sup>36</sup> both consist of two nearly orthogonal antiparallel  $\beta$ -sheets with



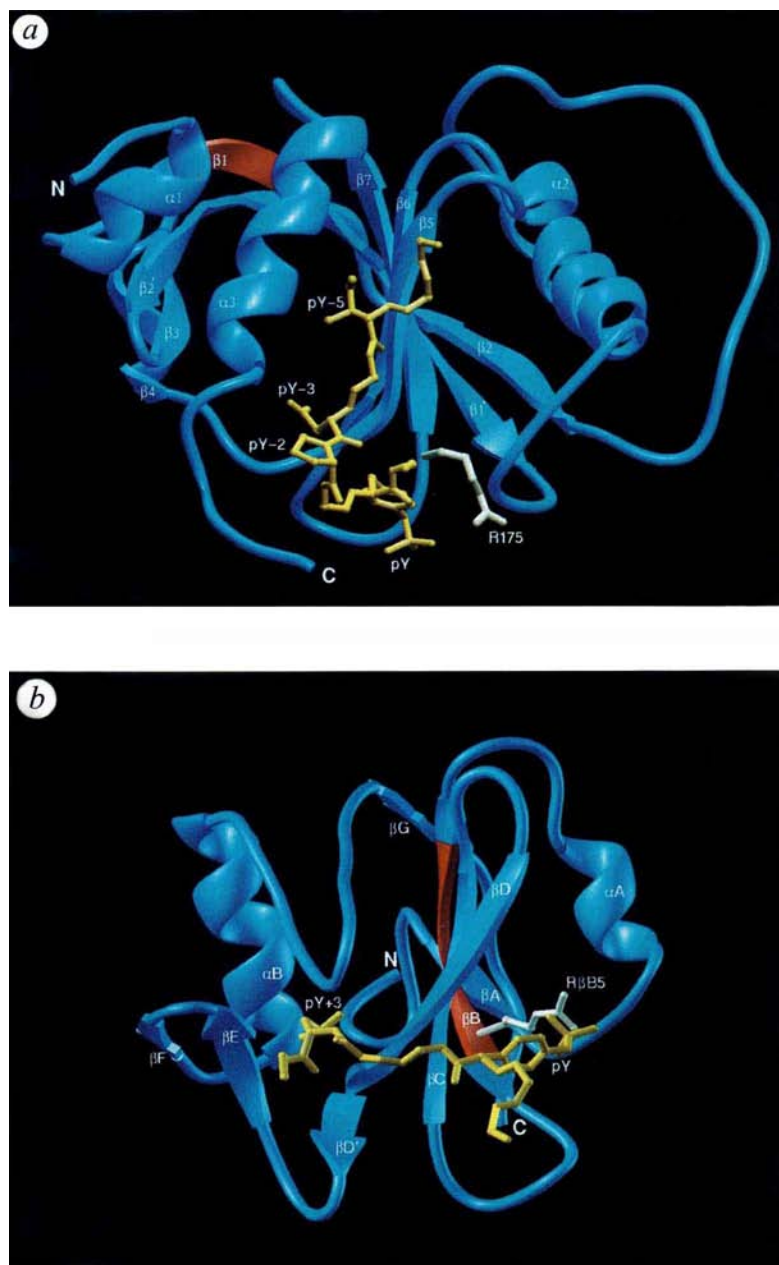


FIG. 4 Ribbon<sup>43</sup> plots depicting *a*, the averaged minimized NMR structure of the PTB domain/phosphopeptide complex compared to *b*, the Shc SH2 domain complexed to a phosphopeptide derived from the  $\zeta$ -chain of the T-cell antigen receptor<sup>34</sup>. Arg  $\beta$ 5 of the SH2 domain and Arg 175 of the PTB domain are shown in white. The F/YLVR sequence is highlighted in red in both domains.

an amphipathic C-terminal  $\alpha$ -helix that interacts with the hydrophobic core of the  $\beta$ -sandwich. The PTB domain is larger than the PH domain, containing an extra N-terminal  $\alpha$ -helix and a large insert between the first and second  $\beta$ -strands. This insert is composed of an  $\alpha$ -helix (residues 72–89) and a large loop (residues 90–108). Another difference is the lack of a tryptophan residue in the C-terminal  $\alpha$ -helix of the PTB domain that is present in all of the PH domains<sup>37</sup>. Despite the structural similarity, there is very little sequence homology between the two domains when the primary sequences of the PTB and PH domains are aligned on the basis of the recently determined three-dimensional structures.

### Binding to acidic phospholipids

PH domains have been shown to bind to the  $\beta\gamma$  subunits of G proteins<sup>38</sup> and acidic phospholipids such as phosphatidylinositol-4-5-bisphosphate (PtdIns(4,5)P<sub>2</sub>)<sup>39</sup>. Because the PTB and PH domains were structurally similar, we examined whether the PTB domain of Shc would bind to phospholipids. As with PH domains<sup>39</sup>, the PTB domain binds to acidic phospholipids but

not to the other lipids tested (Fig. 5*b*). However, in contrast to the PH domains, the PTB domain binds more tightly to PtdIns(4)P (apparent  $K_D$  = 52  $\mu$ M) than to PtdIns(4,5)P<sub>2</sub> (apparent  $K_D$  = 140  $\mu$ M). This binding affinity is similar to that observed for the interaction of lipid vesicles to farnesylated ( $K_D$  = 5–39  $\mu$ M<sup>40</sup>) and myristoylated peptides ( $K_D$  ~ 100  $\mu$ M<sup>40,41</sup>), these moieties being known to play a role in protein-membrane association. The binding to both acidic phospholipids was weaker when the PTB domain was complexed to the TrkA phosphopeptide (Fig. 5*b*). These results suggest that the recruitment of Shc to the membrane could occur in the absence of an interaction with tyrosine-phosphorylated receptors.

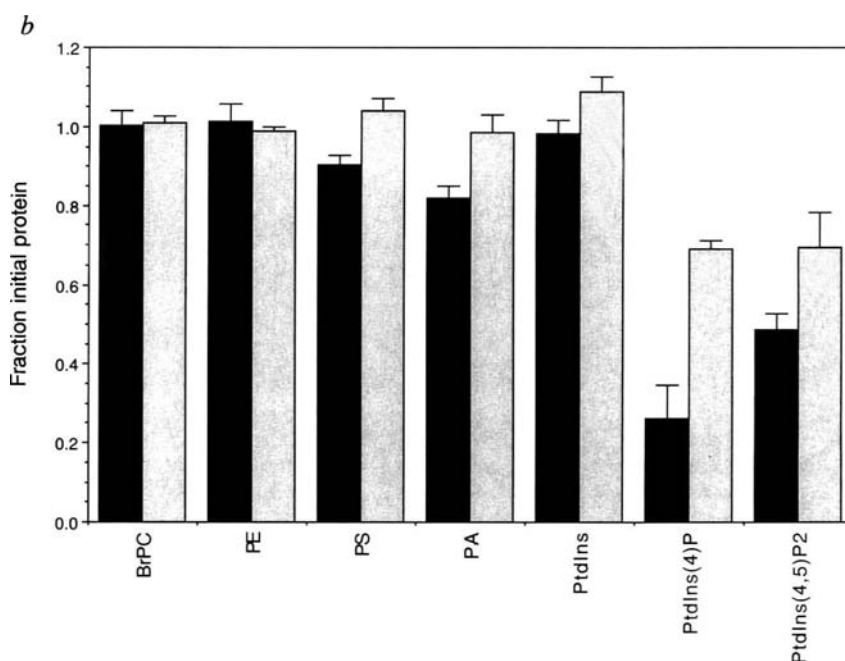
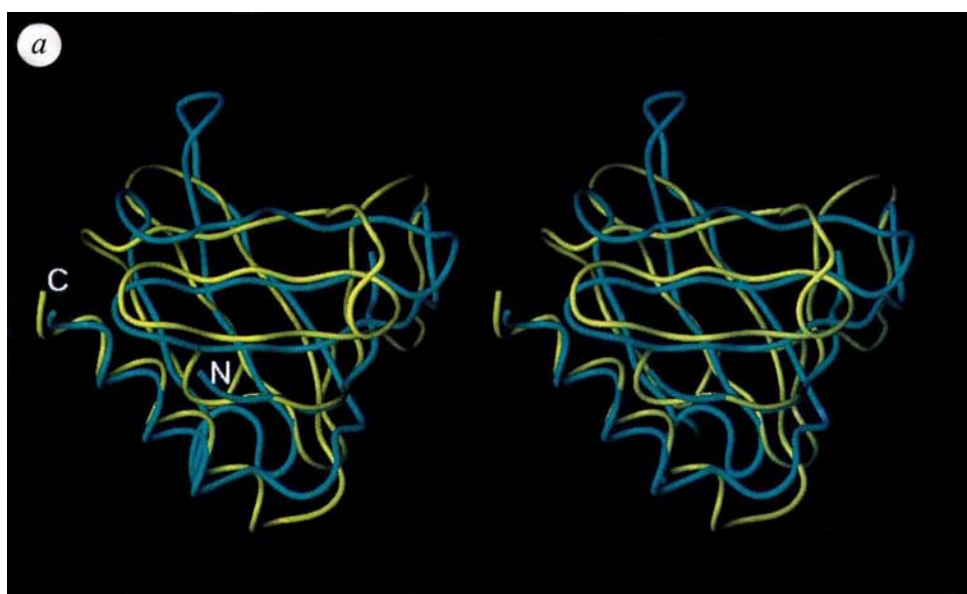
### Discussion

Many signal transduction pathways involve the activation of tyrosine kinases followed by the binding of SH2 domains to tyrosine-phosphorylated proteins. Over a hundred different SH2 domains have been identified, resulting in a significant overlap of binding specificities. The structure of the PTB domain of Shc



FIG. 5 Superposition of the backbone of the NMR structure of the N-terminal PH domain of pleckstrin<sup>47</sup> (yellow) and the PTB domain of Shc (blue). The  $\alpha$  atoms of 65 residues from regions of regular secondary structure were superimposed with an r.m.s.d. of 1.9 Å. **b**, Binding of the PTB domain of Shc in the presence (grey) and absence (black) of the TrkA phosphopeptide to unilamellar vesicles containing 1,2-distearoyl(dibromo)-sn-glycero-3-phosphocholine (BrPC) (Avanti Polar Lipids) and 5% (w/w) of other lipids: phosphoethanolamine (PE), phosphoserine (PS), phosphatidic acid (PA), phosphatidylinositol (PtdIns), phosphatidylinositol-4-phosphate (PtdIns(4)P), and phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P<sub>2</sub>).

**METHODS.** The binding of the PTB domain of Shc to phospholipids was examined with unilamellar vesicles in a centrifugation assay. **1** The vesicles were formed as described previously<sup>48</sup>. Briefly, organic solvent was evaporated from lipid solutions under nitrogen, buffer added, and the lipid suspensions subjected to 5 rapid freeze/thaw cycles. The suspensions were then extruded through 0.1 µm polycarbonate membranes (Costar, Cambridge, MA) ten times. Centrifugation assays were carried out in 50-µl total volume using a Beckman TL-100 ultracentrifuge and a TLA-100 rotor. Protein (25 µl at 0.6 mg ml<sup>-1</sup>) was added to centrifuge tubes. Unilamellar vesicles (25 µl at 10 mg ml<sup>-1</sup> total lipid) of the desired composition were then added, and the mixture incubated 5 min before centrifugation at 100,000 g for 30 min (20 °C). The vesicles were pelleted with this treatment, removing any bound protein from the supernatant. BrPC was used as the background lipid to increase the density of the vesicles and ensure sedimentation at 100,000 g<sup>49</sup>. The protein concentrations of the supernatants (25 µl) was then determined by the BCA protein assay (Pierce, Rockford, IL) and compared with that of protein not exposed to lipid. Association constants ( $K_a$ ) were calculated as suggested<sup>50</sup> assuming a 1:1 complex, from the equation  $K_a = [\text{Shc}]_b / [\text{Shc}]_f [\text{L}]$ , where  $[\text{Shc}]_f$  is the fraction of Shc PTB free in solution after centrifugation,  $[\text{Shc}]_b$  is the amount bound calculated from  $[\text{Shc}]_f$  and the total amount of Shc PTB in the assay, and  $[\text{L}]$  is the concentration of either PtdIns(4)P ( $K_a = 1.9 \times 10^4 \text{ M}^{-1}$ ) or PtdIns(4,5)P<sub>2</sub> ( $K_a = 7.3 \times 10^3 \text{ M}^{-1}$ ) in the outer



leaflet of the unilamellar vesicle (or half the total present). The reciprocal of these numbers can be taken as an apparent dissociation constant<sup>41</sup>.

complexed with the TrkA phosphopeptide shows that binding is accomplished in a very different manner compared to SH2 domains. These structural differences between the PTB and SH2 domains impart markedly different binding specificities for tyrosine-phosphorylated proteins, which is critical for the selection of particular signalling pathways. In this study, we have also identified key residues that are essential for the binding to tyrosine-phosphorylated proteins that may be useful in probing the function of the Shc PTB domain *in vivo*. Moreover, this information may be helpful in identifying other proteins containing the PTB domain.

The PTB domain structure revealed in this study is remarkably similar to the recently determined structures of PH domains. PH domains have been shown to bind to  $\beta\gamma$  subunits of G proteins and acidic phospholipids, and appear to be important for localizing proteins that contain the PH domain to the membrane surface. Many of these proteins associate with phospholipid

membranes, and disruption of this domain may interfere with membrane association. Stimuli through G-protein-coupled receptors, as well as the elevation of intracellular calcium, have recently been shown to lead to tyrosine-phosphorylation of Shc, the formation of the Shc:Grb2:SOS complex, and subsequent Ras activation. However, the manner in which Shc is recruited to the membrane under these conditions (to be in proximity with Ras) is not clear. It is intriguing that the structure of the PTB domain has similarity to PH domains, and that it also binds to acidic phospholipids. It is possible that the PTB domain may help in the localization of Shc to the membrane. If so, such a mechanism may play an important role in linking calcium and G-protein-mediated Ras activation through the use of Shc. □

Received 9 October; accepted 3 November 1995.

1. Giorgetti, S., Pellicci, P. G., Pellicci, G. & Van Obberghen, E. *Eur. J. Biochem.* **223**, 195–202 (1994).

2. Pelicci, G. *et al.* *Cell* **70**, 93–104 (1992).
3. Rozakis-Adcock, M. *et al.* *Nature* **360**, 689–692 (1992).
4. Stephens, R. M. *et al.* *Neuron* **12**, 691–705 (1994).
5. Obermeier, A. *et al.* *EMBO J.* **13**, 1585–1590 (1994).
6. Ravichandran, K. S. & Burakoff, S. J. *J. biol. Chem.* **269**, 1599–1602 (1993).
7. Burns, L. A., Karnitz, L. M., Sutor, S. L. & Abraham, R. T. *J. biol. Chem.* **268**, 17659–17661 (1993).
8. Sato, N., Sakamaki, K., Terada, N., Arai, K. & Miyajima, A. *EMBO J.* **12**, 4181–4189 (1993).
9. Ravichandran, K. S. *et al.* *Science* **262**, 902–905 (1993).
10. Saxton, T., van Oostveen, I., Bowtell, D., Aebersold, R. & Gold, M. J. *Immun.* **153**, 623–636 (1994).
11. Egan, S. E. *Nature* **363**, 45–51 (1993).
12. van Biesen, T. *et al.* *Nature* **376**, 781–784 (1995).
13. Lev, S. *et al.* *Nature* **376**, 737–745 (1995).
14. Kavanaugh, W. M. & Williams, L. T. *Science* **266**, 1862–1865 (1994).
15. Blaikie, P. *et al.* *J. biol. Chem.* **269**, 32031–32034 (1994).
16. Gustafson, T. A., He, W., Craparo, A., Schaub, C. D. & O'Neill, T. J. *Molec. cell. Biol.* **15**, 2500–2508 (1995).
17. Songyang, Z. *et al.* *Molec. cell. Biol.* **14**, 2777–2785 (1994).
18. Songyang, Z., Margolis, B., Chaudhuri, M., Shoelson, S. E. & Cantley, L. C. *J. biol. Chem.* **270**, 14863–14866 (1995).
19. Trüb, T. *et al.* *J. biol. Chem.* **270**, 18205–18208 (1995).
20. Kavanaugh, W. M., Turck, C. W. & Williams, L. T. *Science* **268**, 1177–1179 (1995).
21. van der Geer, P. *et al.* *Curr. Biol.* **5**, 404–412 (1995).
22. Zhou, M.-M. *et al.* *J. biol. Chem.* **270**, 31119–31123 (1995).
23. Dikic, I. *et al.* *J. biol. Chem.* **270**, 15125–15129 (1995).
24. O'Neill, T. J., Craparo, A. & Gustafson, T. A. *Molec. cell. Biol.* **14**, 6433–6442 (1994).
25. Bork, P. & Margolis, B. *Cell* **80**, 693–694 (1995).
26. Sun, X. J. *et al.* *Nature* **377**, 173–177 (1995).
27. Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R. & Kay, L. E. *J. Am. chem. Soc.* **116**, 11655–11666 (1994).
28. Nilges, M., Clore, G. M. & Gronenborn, A. M. *FEBS Lett.* **229**, 317–324 (1988).
29. Kuszewski, J., Nilges, M. & Brünger, A. T. *J. biomolec. NMR* **2**, 33–56 (1992).
30. Willmot, C. M. & Thornton, J. M. *J. molec. Biol.* **203**, 221–232 (1988).
31. Bansal, A. & Gierasch, L. M. *Cell* **67**, 1195–1201 (1991).
32. Eck, M. J., Shoelson, S. E. & Harrison, S. C. *Nature* **362**, 87–91 (1993).
33. Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D. & Kuriyan, J. *Cell* **72**, 779–790 (1993).
34. Zhou, M.-M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **92**, 7784–7788 (1995).
35. Flower, D. R. *FEBS Lett.* **333**, 99–102 (1993).
36. Saraste, M. & Hyvönen, M. *Curr. Opin. struct. Biol.* **5**, 403–408 (1995).
37. Musacchio, A. *et al.* *Trends biochem. Sci.* **18**, 343–348 (1993).
38. Touhara, K. *et al.* *J. biol. Chem.* **269**, 10217–10220 (1994).
39. Harian, J. E., Hajduk, P. J., Yoon, H. S. & Fesik, S. W. *Nature* **371**, 168–170 (1994).
40. Silviu, J. R. & l'Heureux, F. *Biochemistry* **33**, 3014–3022 (1994).
41. Peltzsch, R. M. & McLaughlin, S. *Biochemistry* **32**, 10436–10443 (1993).
42. Brooks, B. R. *et al.* *J. comput. Chem.* **4**, 187–217 (1983).
43. Carson, M. J. *molec. Graph.* **5**, 103–106 (1987).
44. Clore, G. M. & Gronenborn, A. M. *Meth. Enzym.* **239**, 349–363 (1994).
45. Brünger, A. T. *X-PLOR Manual*, version 3.1 (Yale Univ. Press, Cambridge, MA, 1992).
46. Kuboniwa, H., Grzesiek, S., Delaglio, F. & Bax, A. *J. biomolec. NMR* **4**, 871–878 (1994).
47. Yoon, H. S. *et al.* *Nature* **369**, 672–675 (1994).
48. Hope, M. J., Bally, M. B., Webb, G. & Cullis, P. R. *Biochim. biophys. Acta* **812**, 55–65 (1985).
49. Tortorella, D. & London, E. *Analyt. Biochem.* **217**, 176–180 (1994).
50. Rebecchi, M., Peterson, A. & McLaughlin, S. *Biochemistry* **31**, 12742–12747 (1992).

ACKNOWLEDGEMENTS. We thank U. Lorenz for discussions. S.J.B. is supported by an NIH grant, and K.S.R. is supported by a postdoctoral fellowship from the Medical Foundation of the Charles King Trust.

## LETTERS TO NATURE

## Generation of lightning in Jupiter's water cloud

S. Gibbard, E. H. Levy & J. I. Lunine

Lunar and Planetary Laboratory, The University of Arizona, Tucson, Arizona 85721, USA

LIGHTNING is a familiar feature of storms on the Earth, and has also been seen on Jupiter<sup>1–3</sup> and inferred indirectly to occur on Venus and Neptune<sup>4,5</sup>. On Jupiter, lightning may be important as a source of energy to drive chemical reactions in the atmosphere, perhaps determining the abundances of molecules such as CO, HCN and C<sub>2</sub>H<sub>2</sub> (ref. 6). Lightning may be generated in Jupiter's water clouds by a mechanism similar to that which operates in terrestrial thunderstorms<sup>7–9</sup>. Here we investigate the development of lightning by modelling the thunderstorm separation of electrical charge on precipitating ice particles at varying depths in Jupiter's atmosphere. We find that lightning can indeed be generated in the jovian water clouds, and that—in agreement with estimates from the analysis of Voyager images<sup>10</sup>—it is most likely to occur at the 3- or 4-bar pressure level. Our model also predicts that a condensed-water abundance in the range of at least 1–2 g m<sup>-3</sup> is required for lightning to occur in jovian thunderstorms—a prediction that may be tested when the Galileo probe arrives at Jupiter on 7 December 1995.

A lightning discharge occurs when the ambient electric field grows strong enough (the 'breakdown field') that a significant fraction of electrons acquire enough energy in a mean free path ( $\lambda_e$ ) to ionize a molecule in the atmosphere. This leads to a self-perpetuating electron cascade when energy  $E \geq (30 \text{ eV})/e\lambda_e$ , where  $e$  is the electron charge; because of the dependence on  $\lambda_e$ , the breakdown field scales approximately linearly with atmospheric pressure. The actual breakdown field observed at 0.3–0.5 bar in the Earth's atmosphere is typically  $\sim 450,000 \text{ V m}^{-1}$ , lower than the value observed in the laboratory. The difference is thought to be due, in large part, to local field enhancement by ice or water particles<sup>11</sup>.

The global water abundance in Jupiter's troposphere (the atmospheric region of interest) is very uncertain. As discussed in ref. 12, Earth-based spectroscopic data suggest a strong depletion of water relative to its solar abundance, whereas analysis

of Voyager data indicates values exceeding solar abundance (0.13% by number relative to the primary constituent, molecular hydrogen). Other estimates from the inferred mass distribution of Jupiter suggest water-abundance values of 2–5 times solar and up to 10 times solar from Shoemaker-Levy data<sup>12</sup>. One possible resolution of these disparate indications might lie in a heterogeneous distribution of water associated with convective storms<sup>13</sup>. Our model of jovian thunderstorm electrification does not assume that the condensed water abundance in the lightning-producing region must be the same as the global water abundance. We are accordingly able to test the sensitivity of the lightning model to variations in the amount of condensed water. Our model also incorporates the levitation of particles due to the electrostatic force, which plays an important role in determining the charging timescale at higher pressures.

In a recent study, Yair *et al.*<sup>14</sup> analysed the detailed microphysics of the electrical development of a jovian water-cloud thunderstorm. This study was based on a two-dimensional, axially symmetric, convection-precipitation model of the lightning-producing thundercloud<sup>15</sup>, and non-inductive charge transfer between ice and graupel particles. Yair *et al.* modelled the development of the large-scale electric field, with atmospheric water abundance comparable to that of the Sun, and found that the electric field grew to exceed the breakdown value after about 1.5 hours into the storm, with the electric field maximum occurring at a pressure level of  $\sim 2.5$  bar. The model of ref. 14 provides a detailed picture of the growth, electrification, and eventual disintegration of a jovian thundercloud.

The present Letter reports a less detailed, but complementary, study of thundercloud development, investigating the possibility of lightning occurrence at different levels in the jovian atmosphere. The calculations reported here are based on a one-dimensional convection-precipitation model, with parametrized edge entrainment—an approach that captures the essence of the two-dimensional effects, and the main characteristics of the upwelling cloud, well enough for the purpose of exploring the development of lightning-producing electric fields. We report a set of cloud-electrification investigations, also assuming non-inductive charging mechanisms, referenced to several different pressure levels in the jovian atmosphere. Our results are consistent with those of ref. 14, encompassing the possibility of lightning production at 2.5 bar pressure. Moreover, our pressure-parametrized study places specific limits on the broader altitude band of the jovian

atmosphere in which lightning can occur, and outside of which lightning would not be expected.

Our model assumes collisional charging of ice particles, as observed in laboratory experiments<sup>16,17</sup>, and gravitational separation of oppositely charged large and small particles due to their differential fall velocities. At temperatures below  $-15^{\circ}\text{C}$ , large particles are usually observed to acquire a negative charge in collisions with smaller ones<sup>17</sup>. We use a spectrum of particle sizes grown by sticking collisions of  $10\text{-}\mu\text{m}$  cloud particles to a maximum radius of  $2\text{--}4\text{ mm}$ . Differential fall velocities are determined as a function of time by calculating the total force on each particle due to gravity, atmospheric drag, and the electric force. The amount of charge transferred per collision was determined by using a piecewise linear fit to the experimental data of Keith and Saunders<sup>18</sup>, which involved ice-crystal interactions with a riming target. Details of particle growth and charging will be given elsewhere (S.G., E.H.L. and J.I.L., manuscript in preparation).

We assume that lightning is generated in association with 'moist' convection, which refers to the role that condensation of (in this case) water plays in modifying the vertical motions associated with convective transport of heat through the atmosphere, especially as latent heat warms the gas when water condenses in a rising plume. To simulate moist convection, we employ a simple one-dimensional model known as the "entraining plume". In this model, a buoyant column of hydrogen gas and condensing water draws in gas from the atmosphere that surrounds the column during the uplift. This model was first applied to Jupiter by Stoker<sup>19</sup>, and, although it makes a number of simplifications, provides a convenient representation of the lightning-generating environment without an excessive number of adjustable parameters. The model has been extended with more accurate computation of the condensate loading using a terrestrial convective formalism<sup>20</sup> and corrected for jovian conditions by comparison with axisymmetric models of moist convection<sup>15</sup>. The entraining-plume model assumes that a source of uplift is available to raise moist parcels to their level of buoyancy, and then computes the upward velocity of the buoyant plume as a function of altitude.

We choose plume diameters of 5 and 10 km, just above the size range for which dry air entrainment inhibits vertical convection, and consistent with other size considerations<sup>12,15</sup>. Although

convective storm systems on Jupiter may have greater areal extent, one-dimensional models cannot directly account for the structure of such systems, which are an ensemble of upward and downward moving plumes. Choosing a size scale at the lower end of the plausible range indirectly accounts for this complexity in the structure of moist convective clouds.

We find, for a range of global background water abundances from solar to several times this value, and relative humidities in the environment of  $\leq 70\%$ , that self-sustaining buoyant moist convection is initiated in the 4- to 6-bar region, even though the cloud base itself may be lower. Peak upward velocities in the plume range from  $20$  to  $100\text{ m s}^{-1}$ , although the upper value is almost certainly excessively high, based on axisymmetric models<sup>15</sup>. Peak velocities of several tens of metres per second, as are achieved in our model for environmental relative humidities of  $< 50\%$ , appear reasonable. Peak condensate loads of several grams per cubic metre are predicted, also in agreement with the two-dimensional models<sup>15</sup>. Buoyancy in these clouds is lost at  $\sim 0.5$  bar.

We also applied our model to the ammonium hydrosulphide cloud; vapour pressures of reactions between hydrogen sulphide and ammonia that form ammonium hydrosulphide condensates are available<sup>21</sup>. Such clouds have their base around the 2- to 3-bar region, for sulphur and nitrogen abundances between 1 and 4 times solar, but are negatively or near-neutrally buoyant throughout their vertical extent. Although the model predicts velocities of up to  $30\text{ m s}^{-1}$  over restricted altitude regimes (because, unlike in the water cloud, the condensate load is very low), the near-neutral buoyancy suggests that, under real jovian conditions, substantial vertical moist convection might not occur at all in these clouds. Thus unlike the water clouds in the 6- to 1-bar region, the hydrosulphide clouds seem to be poor candidates for lightning.

We calculated water cloud models for 5, 4, 3 and 2 bar. Deeper than  $\sim 5$  bar the temperature should be above the freezing point of water. In the absence of an ice phase, charge separation is inefficient. Thus a significant amount of jovian lightning at pressures above 5 bar is unlikely. (Observations show that terrestrial lightning is only very infrequently generated from entirely liquid clouds<sup>22</sup>.) At pressures substantially below 2.7 bar the temperature remains below  $-45^{\circ}\text{C}$ , and charge separation begins to be inhibited by the lack of a liquid phase and by the

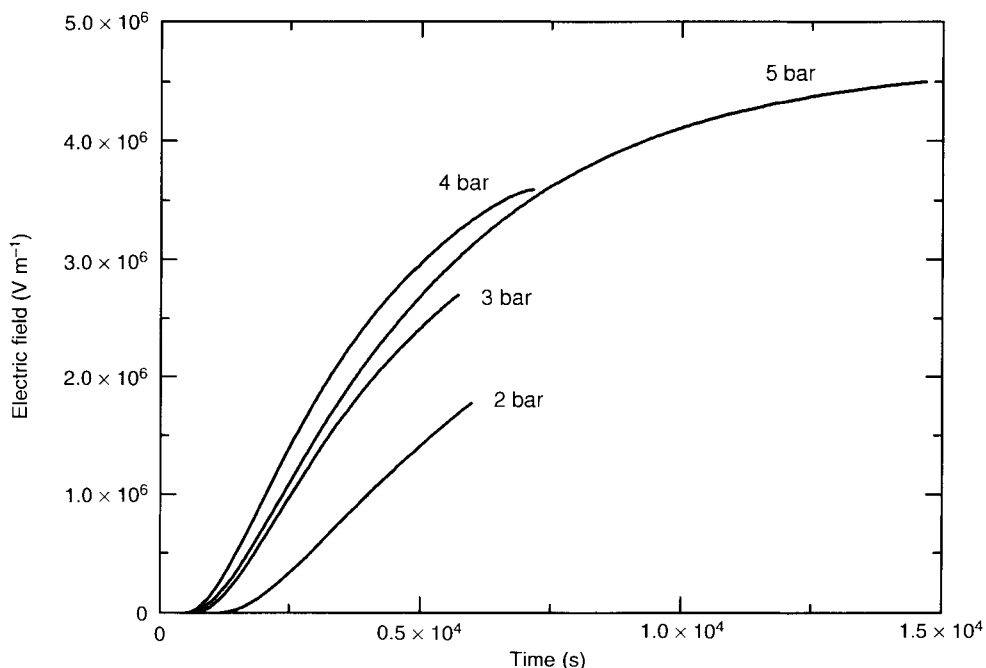


FIG. 1 Temporal development of the large-scale electric field in a jovian thunderstorm. The electric field is shown as a function of time for four pressure levels in the atmosphere, with the standard baseline water abundances indicated in the text. Each curve terminates at the break-down field corresponding to the respective pressure.



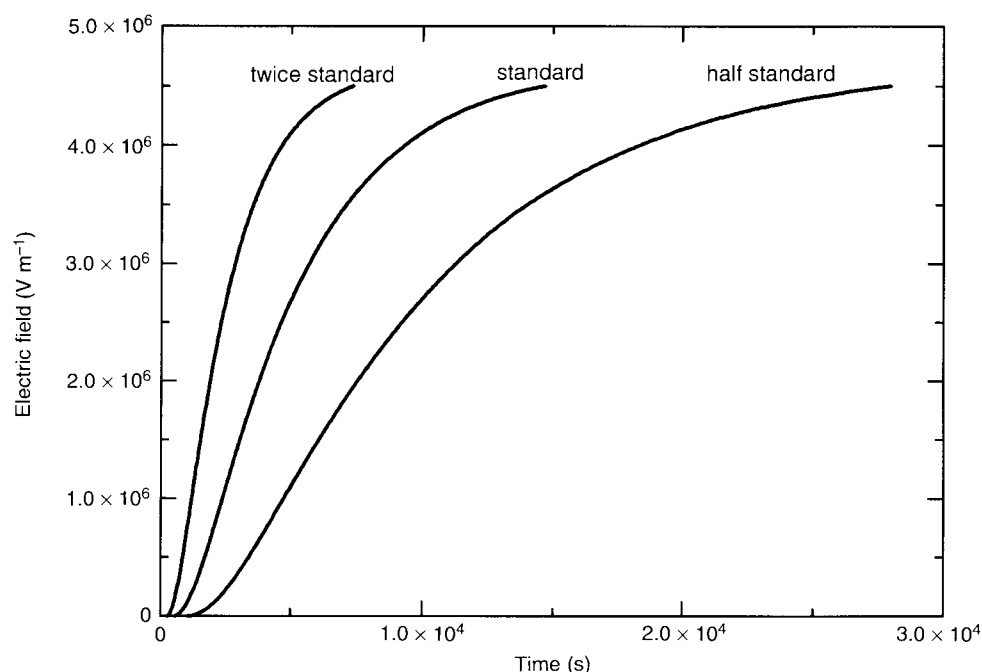


FIG. 2 The temporal development of the large-scale electric field, at a pressure level of 5 bar, for water abundances below, at, and above the standard baseline values. As in Fig. 1, each curve terminates at the breakdown field intensity.

low electrical conductivity of the cold ice particles<sup>11</sup>. Also, the water abundance is low at such high altitudes.

Results of our calculations are shown in Figs 1 and 2. The background water abundance, outside the storms, is taken to be four times the solar value, consistent with the discussion above. The moist convective model yields values for the condensed water abundance in the cloud of one to several grams per m<sup>3</sup>. Although the actual condensed water abundance is sensitive to various microphysical parameters that are hard to characterize for Jupiter (for example, the presence and properties of condensation nuclei, and details of updraft and downdraft geometries), we adopted a baseline condensed water profile of 1 g m<sup>-3</sup> at 2-bar pressure level, 2 g m<sup>-3</sup> at 3 bar and 3 g m<sup>-3</sup> for 4–5 bar. These values are consistent with typical results of the entraining plume model, as well as with more complex axisymmetric models<sup>15</sup>. Figure 1 shows the growth of the large-scale electric field with time for the baseline condensed water abundance at several pressure levels. Figure 2 shows the electric field growth at 5 bar for several different values of the water abundance: the baseline values cited, twice those values and half those values.

For the baseline water abundance, the electric field strength needed to produce lightning can be achieved in ~6,000 s at the 2- or 3-bar level, 7,000 s at the 4-bar level, but requires almost 15,000 s at the 5-bar level. Electrostatic levitation is an important effect at the higher breakdown field values associated with higher pressures—causing oppositely-charged small and large particles to move together at a common velocity, thus inhibiting large-scale charge separation. At twice the baseline water abundance, charging is considerably more rapid. At half the baseline abundance, the charging is so slow that particles fall a large fraction of an atmospheric scale-height (~35 km) before breakdown is achieved. In this latter case, lightning is inhibited because particles rain out before sufficient charge is separated.

Jovian lightning seems most likely to occur at the 3- or 4-bar level, where the temperature allows effective charge separation and the charging timescale is short enough that breakdown-field levels can be achieved before the particles completely rain out. At pressures approaching 5 bar, electrostatic levitation of the precipitating particles slows electrification to the point that the particles will rain out before breakdown is achieved, unless the water abundance is substantially higher than the baseline

(middle) value taken here. At altitudes approaching the 2-bar level, the low ambient temperature (~200 K) is expected to reduce the efficiency of charge transfer<sup>23</sup> below what was used in these calculations. The combination of that effect and the low water abundance inhibits lightning generation at such high altitude. Moreover, if the water abundance is significantly lower than the baseline values used here, the timescales would become too long even at the 3- and 4-bar levels, indicating that lightning would not occur at much lower condensed water abundances in jovian storms.

The descent of the Galileo probe in December 1995 will improve our knowledge of the jovian water abundance and cloud structure. Our model calculations provide a way to determine a minimum condensed water abundance in storm regions, because we have found that the presence of lightning implies the presence of a certain minimum of condensed water. This could provide an important comparison with results of the nephelometer and mass-spectrometer measurements. Most crucial will be Galileo determinations of the water abundance, both at high vertical resolution from the probe's neutral mass spectrometer and over the disk of Jupiter from the orbiter's near-infrared mapping spectrometer. A deep water abundance exceeding the mapping value, and a generally heterogeneous water distribution in the cloud-forming region, would lend support to the conditions of moist convection assumed here. This in turn provides a natural physical context for the lightning discharges detected by Voyager, and those which Galileo may detect using several experiments. Assembling a consistent picture of cloud convection and lightning generation is essential if we are to understand how the Solar System's largest planet transports heat through its atmosphere and gain a full understanding of the energetic process that drive the planet's atmospheric chemistry. □

Received 15 May; accepted 17 November 1995.

1. Cook, A. F., Duxbury, T. C. & Hunt, G. E. *Nature* **280**, 794 (1979).
2. Smith, B. A. *et al.* *Science* **204**, 951–972 (1979).
3. Scarf, F. L., Gurnett, D. A. & Kurth, W. S. *Science* **204**, 991–995 (1979).
4. Rinnert, K. J. *geophys. Res.* **90**, 6225–6237 (1985).
5. Russell, C. T. A. *Rev. Earth Planet. Sci.* **21**, 43–87 (1993).
6. Podolak, M. & Bar-Nun, A. *Icarus* **75**, 566–570 (1988).
7. Borucki, W. J. & Magalhães, J. A. *Icarus* **96**, 1–14 (1992).
8. Levin, Z., Borucki, W. J. & Toon, O. B. *Icarus* **56**, 80–115 (1983).
9. Uman, M. *The Lightning Discharge* (Academic, Orlando, FL, 1987).

10. Borucki, W. J. & Williams, M. A. *J. geophys. Res.* **91**, 9893–9903 (1986).
11. Gross, G. W. *J. geophys. Res.* **87**, 7170–7178 (1982).
12. Ingersoll, A. P. & Kanamoori, H. *Nature* **374**, 706–708 (1995).
13. Lunine, J. I. & Henten, D. M. *Icarus* **60**, 566–570 (1987).
14. Yair, Y., Levin, Z. & Tzivion, S. *Icarus* **115**, 421–434 (1995).
15. Yair, Y., Levin, Z. & Tzivion, S. *Icarus* **114**, 278–299 (1995).
16. Caranti, J. M., Illingworth, A. J. & Marsh, S. J. *J. geophys. Res.* **90**, 6041–6046 (1985).
17. Norville, K., Baker, M. & Latham, J. J. *J. geophys. Res.* **96**, 7463–7481 (1991).
18. Keith, W. D. & Saunders, C. P. R. *J. geophys. Res.* **94**, 13103–13106 (1989).
19. Stoker, C. R. *Icarus* **67**, 106–125 (1986).
20. Yanai, M., Esbensen, S. & Chu, J. J. *J. atmos. Sci.* **30**, 611–627 (1973).
21. Weast, R. C. *CRC Handbook of Chemistry and Physics* (Chemical Rubber Company, Cleveland, OH, 1967).
22. Takahashi, T. *J. atmos. Sci.* **32**, 123–141 (1975).
23. Illingworth, A. J. *J. geophys. Res.* **90**, 6026–6032 (1988).

## Evidence from molecular dynamics simulations for non-metallic behaviour of solid hydrogen above 160 GPa

John S. Tse & Dennis D. Klug

Steacie Institute for Molecular Sciences, National Research Council of Canada, Ottawa, Ontario K1A 0R6, Canada

THE behaviour of molecular hydrogen at high pressures has implications for the interiors of the giant planets, which consist mainly of hydrogen. In particular, the question of whether solid hydrogen becomes metallic under these conditions has been much debated<sup>1,9</sup>, in part because the structure that molecular hydrogen adopts at high pressure is not known. Here we report the results of first-principles molecular dynamics simulations of solid hydrogen at pressures up to 270 GPa. We find that at 77 K, hydrogen exists as a stable, orientationally disordered phase up to 60 GPa, consistent with experimental results<sup>1,10</sup>. As the pressure is raised, a gradual transformation to an ordered orthorhombic structure begins at 160 GPa, and by 260 GPa the solid becomes semi-conducting, with an indirect band gap of 1.4 eV. The calculated vibrational density of states of this phase is consistent with infrared and Raman spectra measured up to 160 GPa (ref. 11). Although limitations on the simulation time and size may result in an overestimate of the absolute pressure, our calculations show that solid hydrogen does not become metallic, even at pressures approaching 260 GPa.

Our first-principles molecular dynamics calculations are based on local density-functional theory and pseudopotential plane-wave approximations<sup>12</sup>. A principal advantage of this approach is that possible structures can be advantaged without making assumptions about the space-group symmetry. A previous study<sup>13</sup> has shown that introduction of zero-point vibrations (ZPVs) in an *ad hoc* manner by adding the contribution of selected phonon modes may affect the stability of competing low-pressure phases. However, for orientationally ordered phases at high pressures, this effect is not likely to be severe<sup>9</sup>. This is strongly supported, for example, by the observation that transition pressures are very similar for hydrogen and deuterium<sup>1</sup>. A more recent theoretical study<sup>9</sup> on solid hydrogen showed that the molecular orientation and the transition pressures to various high-pressure phases are not affected by the ZPVs. We have not taken ZPVs into account here. A rigorous treatment taking quantum effects into account can be obtained with the path-integral technique<sup>14</sup>, but this method does not provide dynamical information.

Even at low pressures, where ZPVs are expected to be most important, the calculated equation of state agrees well with the extrapolated experimental data<sup>1</sup>. Below 60 GPa, the hydrogen molecules are rotating almost freely about hexagonal lattice sites. As the pressure is increased, the rotation ceases and an orienta-

tionally ordered structure is formed with the molecules aligned in layers parallel to the *a-b* plane. This structural transformation is gradual; it begins at about 160 GPa and is complete by 260 GPa. The large transition region is an artefact associated with the use of a small simulation cell and short annealing time, and therefore, the experimental transition is expected to be sharper and to occur at pressures lower than 260 GPa. The variation of intramolecular hydrogen bond lengths at several pressures is shown in Fig. 1. At 15 GPa, the bond length distribution is sharply centred at 1.422 *a*<sub>0</sub>, only 0.005 *a*<sub>0</sub> shorter than the calculated gas-phase value (1 *a*<sub>0</sub> = 0.529 Å). Intermolecular interactions are obviously insignificant and the barrier for molecular rotation is very small. At 100 GPa, the large-amplitude rotational motions are restricted to the *a-b* plane and the hydrogen bond lengths are shortened to 1.390 *a*<sub>0</sub>. This ordered structure remains stable up to 150 GPa at which point the hydrogen bond length shrinks further to 1.370 *a*<sub>0</sub>. On further pressu-

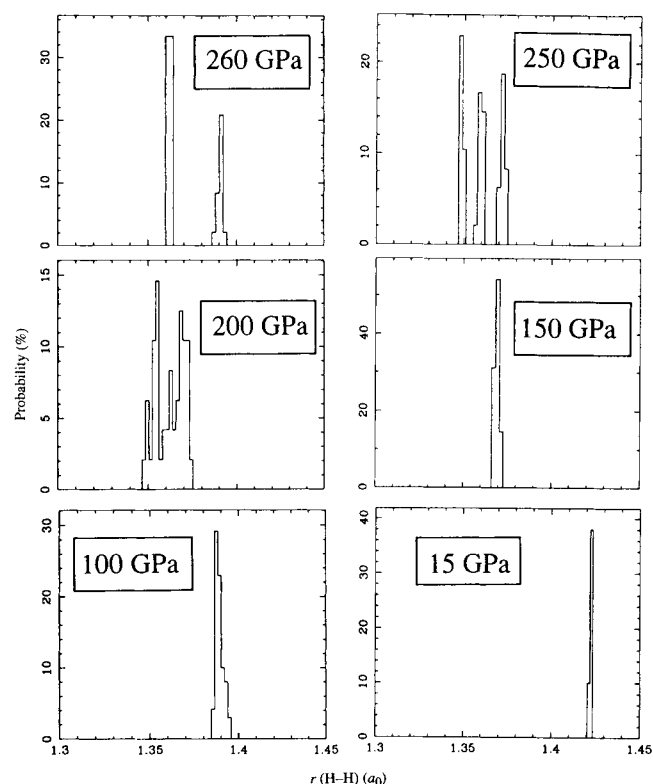


FIG. 1 The normalized distribution of intramolecular hydrogen bond lengths at the pressures indicated. Calculations were performed on 14 densities at 77 K (corresponding to pressures in the range 15–270 GPa) with a supercell consisting of 96 hydrogen atoms with the centre of the hydrogen molecules initially located at the hexagonal close-packed sites as proposed by Kaxiras *et al.*<sup>7</sup>. It was found that a smaller 64-atom cell was not adequate for the convergence in energy and molecular properties<sup>22</sup>. For most calculations, the *c/a* ratio of the hexagonal cell was maintained at the ideal value of 1.633 while the atom positions were allowed to relax. Additional optimization at several selected pressures show that this value is almost optimal. An ultrasoft pseudopotential<sup>23</sup> with an energy cutoff of 20 rydberg for the hydrogen was used. This pseudopotential yields an H<sub>2</sub> equilibrium bond length of 1.427 *a*<sub>0</sub> and a vibrational frequency of 4,034 cm<sup>-1</sup> which compare favourably with the experimental values<sup>24</sup> of 1.400 *a*<sub>0</sub> and 4,161 cm<sup>-1</sup>, respectively. The Perdew–Zunger<sup>25</sup> parametrization of the Monte Carlo results of Ceperley and Alder for the exchange-correlation potential in a homogeneous electron gas corrected for the nonlocal contribution was employed<sup>26,27</sup>. The  $\Gamma$  (*k*=0) point (see Fig. 4) is used for Brillouin zone sampling. At each pressure, the equilibrium structure was obtained by simulated annealing and the trajectories were then collected for at least 20,000 time steps of 7.257 × 10<sup>-17</sup> s after thermal equilibration where the structural and dynamical information was extracted.

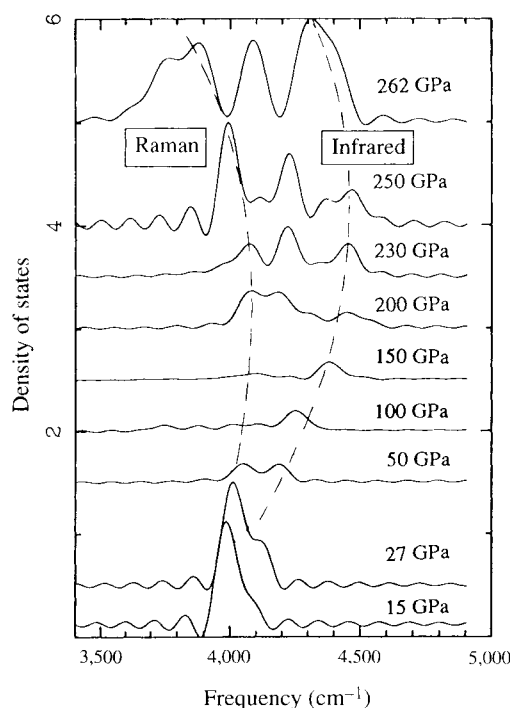


FIG. 2 The density of vibrational states of solid hydrogen calculated as a function of pressure. The lines are drawn to guide the eye. The intensities are in arbitrary units.

rization, the bond lengths are distributed from  $1.35 a_0$  to  $1.37 a_0$ , suggesting the coexistence of two phases. At about 260 GPa the transformation to a new ordered structure having two distinct intramolecular hydrogen bond lengths, with a short ( $1.35 a_0$ ) to long ( $1.37 a_0$ ) population ratio of exactly 2:1 is evident. Extensive rearrangements of the electronic structure are clearly associated with the structural changes.

The vibrational density of states (VDOS), comprised of all Raman and infrared vibrations, can be evaluated from the Fourier transformation of the atom velocity autocorrelation function. The calculated H-H stretching vibrations in the range  $3,500\text{--}4,500\text{ cm}^{-1}$  are depicted in Fig. 2. At low pressure, the VDOS consist of a band at  $\sim 4,000\text{ cm}^{-1}$  with a distinct higher-frequency shoulder. The vibrational frequency is only  $40\text{ cm}^{-1}$  lower than the gas-phase value, a clear indication of the weak interactions in the solid. When the pressure is raised the stretching band splits into two components and gradually shifts to higher frequencies, reaching a maximum of  $4,400\text{ cm}^{-1}$ . This trend continues until  $\sim 150\text{ GPa}$ , where a new band starts to emerge at a lower frequency of  $\sim 4,100\text{ cm}^{-1}$ . At  $200\text{ GPa}$  the VDOS of the stretching vibration region consists of three distinct bands at  $4,000$ ,  $4,150$  and  $4,450\text{ cm}^{-1}$ , and the frequencies shift to lower energy as the pressure is further increased. The higher-energy modes may be correlated with the infrared vibrations, and the lower-energy bands with Raman vibrations. The calculated trend shows qualitative agreement with experimental findings<sup>15–17</sup>. In the experiments, the frequency of the Raman band was found to increase slightly with pressure, and at  $150\text{ GPa}$  a new band at lower frequency appears<sup>23</sup> and the original band disappears. Very similar behaviour was found in the infrared studies<sup>18</sup>, except that the absorptivity of the new absorption band increases dramatically.

The abrupt increase in the infrared absorptivity has been rationalized by a charge-transfer mechanism<sup>11</sup>. Several structural models have been recently suggested<sup>17,19</sup>, but a consistent picture of how the molecules rearrange to facilitate this kind of interaction from the low-pressure structure is still lacking. Our first-

principles molecular dynamics calculations show a displacive transformation to an orthorhombic structure at  $160\text{--}260\text{ GPa}$  (Fig. 3). This resembles suggested high-pressure forms<sup>1,7</sup> as the average positions of the centres of the hydrogen molecules displace only slightly from the ideal hexagonal sites. The unique feature of this new structure is the grouping of three hydrogen molecules originally in the layers parallel to the  $a\text{--}b$  plane, alternately shifted along the  $c$  axis with concomitant rotation of the molecules out of the plane. The unit cell is composed of four components that are related by symmetry. Each component consists of two equivalent hydrogen molecules with a short bond length of  $1.35 a_0$  situated on both ends of the third hydrogen molecule with a longer bond length of  $1.37 a_0$ . The closest distance between the short- and long-bonded hydrogen molecules is  $2.35 a_0$ . This special conformation and grouping of the hydrogen molecules has a profound effect on the electronic properties. Band-structure calculations show that there is small overlap between the hydrogen molecules. In particular, there is a slight accumulation of electronic charge in the longer-bonded hydrogen molecules with electron depletion in the shorter ones. These charge-transfer interactions are dependent on the unique conformation and can be explained qualitatively with the aid of the molecular orbital (MO) diagram shown in Fig. 3. Six sets of orbitals can be derived from the linear combination of the three bonding and anti-bonding molecular orbitals from each molecule<sup>20,21</sup>. The relative energy ordering of the MOs can be assigned in accordance with the number of nodal planes. At high pressures, when the molecules are pushed closer together, the

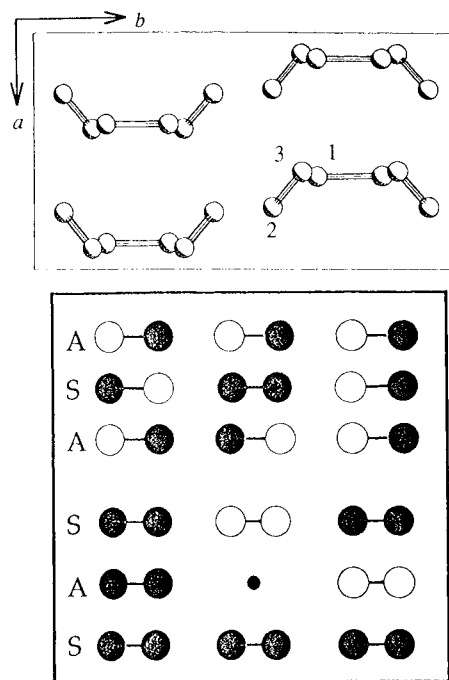


FIG. 3 Top, the structure of the orthorhombic phase of solid hydrogen optimized at  $260\text{ GPa}$  ( $Pnma$  with  $a=2.767\text{ Å}$ ,  $b=4.971\text{ Å}$  and  $c=2.870\text{ Å}$ ). There are only three crystallographically distinct hydrogen atoms as labelled in the diagram. The pressure was estimated from the theoretical equation of states which closely corresponds to the equation of state proposed by Hemley *et al.*<sup>28</sup>. Bottom, a schematic molecular orbital diagram for the interacting orbitals in a three-molecule subunit of hydrogen. The open and shaded circles represent the phase of the component orbitals. The dot represents a nodal plane. The structure is optimized by varying the  $c/a$  ratio where a shallow energy minimum is located at  $1.67$ . The cell parameters are found to be strongly correlated and the accurate determination of the minimum-energy structure require the calculation of the stress tensor<sup>16,17</sup> and a more elaborate  $k$ -point sampling scheme.



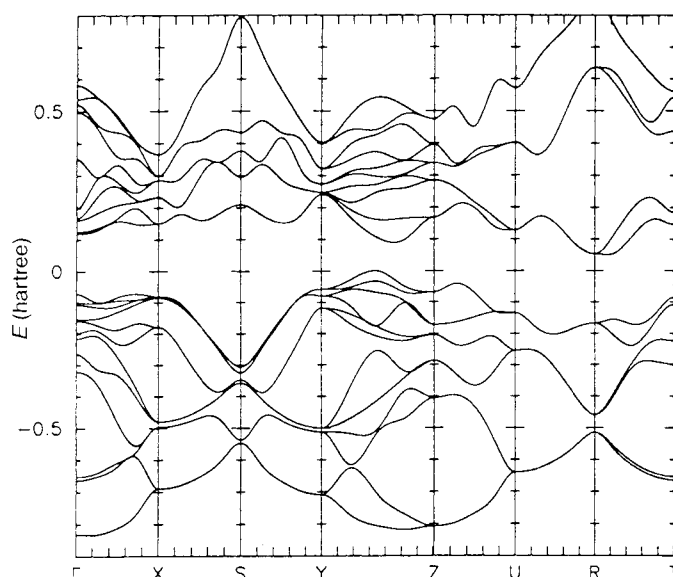


FIG. 4 The calculated electronic band structure of the *Pnma* structure of solid hydrogen at 260 GPa. The calculations were performed with a numerical basis set augmented with localized *p*-type orbitals. A total of 163 *k*-points were used in the Brillouin zone integration<sup>29</sup>.

first symmetric (S) orbital in the bonding set are stabilized but the effect on the asymmetric (A) 'non-bonding' orbital is small. However, the highest occupied symmetric orbital will be destabilized and the orbital energy shifted upwards due to out-of-phase interactions between the component orbitals. In the empty orbital set, the first unoccupied asymmetric orbital will decrease in energy because of the favourable in-phase interactions between the molecules. The other two remaining empty orbitals are expected to shift to higher energy owing to increased anti-bonding interactions. The empty asymmetric anti-bonding orbital that is shifted to lower energy can then efficiently interact with the occupied asymmetric 'non-bonding' orbital. The net effect is the destabilization of the central hydrogen molecule due to the increase in the anti-bonding character resulting in a donation of charge density from the two peripheral hydrogen molecules to the central one, forming a weak charge-transfer complex. Consequently, the bond of the central hydrogen lengthens relative to the other two hydrogens which are equivalent by symmetry. The small charge transfer creates a dipole moment in the shorter-bonded hydrogen molecules leading to the enhancement in the infrared absorptivity. In contrast, the longer-bonded hydrogen molecule becomes slightly negatively charged. Owing to crystal reflection symmetry, this molecule cannot possess a permanent dipole moment and hence the stretching vibration will only be Raman-active and the stretching frequency is expected to be lower. The observation of a new Raman band at low frequency at 160 GPa can be attributed to this effect.

We calculate an indirect band gap of 1.4 eV at 260 GPa for this structure (Fig. 4). Thus, even at this pressure (which is considerably higher than experiments have yet reached), we predict that solid hydrogen will remain non-metallic. The Fermi level is located along the Z→Y line and the bottom of the conduction band is at the R point. (Here the letters represent symmetry points on the surface of the first Brillouin zone. At the Y point, the Bloch crystal orbital is destabilized owing to the out-of-phase interactions of the highest occupied symmetric MOs of the molecular fragments whereas at the R point, the crystal orbital is stabilized by the out-of-phase combination of the first empty asymmetric MOs of the molecular fragments. Eventual metallization through the closure of the band gap is likely to occur at these symmetry points at a higher pressure. □

Received 21 March; accepted 24 October 1995.

1. Mao, H.-K. & Hemley, R. J. *Rev. Mod. Phys.* **66**, 671–692 (1994).
2. Chakravarty, S., Rose, J. H., Wood, D. & Ashcroft, N. W. *Phys. Rev.* **B24**, 1624–1635 (1981).
3. Min, B. I., Jansen, H. J. F. & Freeman, A. J. *Phys. Rev.* **B33**, 6383–6390 (1986).
4. Barbee, T. W., Garcia, A., Cohen, M. L. & Martins, J. L. *Phys. Rev. Lett.* **62**, 1150–1153 (1989).
5. Chacham, H. & Louie, S. G. *Phys. Rev. Lett.* **66**, 64–67 (1991).
6. Garcia, A., Barbee, T. W., Cohen, M. L. & Silvera, I. F. *Europhys. Lett.* **13**, 355–360 (1990).
7. Kaxiras, E., Broughton, J. & Hemley, R. J. *Phys. Rev. Lett.* **67**, 1138–1141 (1991).
8. Nagara, H. & Nakahara, T. *Phys. Rev. Lett.* **68**, 2468–2471 (1992).
9. Natoli, V., Martin, R. M. & Ceperley, D. *Phys. Rev. Lett.* **74**, 1601–1604 (1995).
10. Mao, H. K. et al. *Science* **239**, 1131–1134 (1988).
11. Hemley, R. J., Soos, Z. G., Hanfland, M. & Mao, H.-K. *Nature* **369**, 384–387 (1994).
12. Car, R. & Parrinello, M. *Phys. Rev. Lett.* **55**, 2471–2474 (1985).
13. Surh, M. P., Barbee, T. W. & Mailhot, C. *Phys. Rev. Lett.* **70**, 4090–4093 (1993).
14. Marx, D. & Parrinello, M. *Z. Phys.* **B95**, 143–144 (1994).
15. Sharma, S. K., Mao, H. K. & Bell, P. M. *Phys. Rev. Lett.* **44**, 886–888 (1980).
16. Hemley, R. J. & Mao, H. K. *Phys. Rev. Lett.* **61**, 857–860 (1988).
17. Hanfland, M., Hemley, R. J., Mao, H. K. & Williams, G. P. *Phys. Rev. Lett.* **69**, 1129–1132 (1992).
18. Hanfland, M., Hemley, R. J. & Mao, H. K. *Phys. Rev. Lett.* **70**, 3760–3763 (1993).
19. Kaxiras, E. & Broughton, J. *Comput. Mater. Sci.* **3**, 368–376 (1995).
20. Hoffmann, R. *Solids and Surfaces* (VCH, New York, 1988).
21. Albright, T. A., Burdett, J. K. & Whangbo, M.-H. in *Orbital Interactions in Chemistry* (Wiley, New York, 1985).
22. Hohl, D. et al. *Phys. Rev. Lett.* **71**, 541–544 (1993).
23. Vanderbilt, D. *Phys. Rev.* **B41**, 7892–7895 (1990).
24. Wijngaarden, R. J., Lagendijk, A. & Silvera, I. F. *Phys. Rev.* **B26**, 4957–4961 (1982).
25. Perdew, J. P. & Zunger, A. *Phys. Rev.* **B23**, 5048–5078 (1981).
26. Perdew, J. P. *Phys. Rev.* **B33**, 8822–8824 (1986).
27. Becke, A. D. *Phys. Rev.* **B38**, 3098–3100 (1988).
28. Hemley, R. J. et al. *Phys. Rev.* **B42**, 6458–6470 (1990).
29. te Velde, G. & Baerends, E. J. *Phys. Rev.* **B44**, 7888–7903 (1991).

## Experimental confirmation of universality for a phase transition in two dimensions

C. H. Back, Ch. Wüsch, A. Vaterlaus,  
U. Ramsperger, U. Maler & D. Pescia

Laboratorium für Festkörperphysik der ETH Zürich,  
CH-8093 Zürich, Switzerland

WHEN a system is poised at a critical point between two macroscopic phases, it exhibits dynamical structures on all available spatial scales, even though the underlying microscopic interactions tend to have a characteristic length scale. According to the universality hypothesis<sup>1,2</sup>, diverse physical systems that share the same essential symmetry properties will exhibit the same physical behaviour close to their critical points<sup>1,3–5</sup>; if this is so, even highly idealized models can be used to describe real systems accurately. Here we report experimental confirmation that the scaling behaviour of thermodynamic variables predicted by the universality hypothesis holds over 18 orders of magnitude. We show that the equation of state of a two-dimensional system (an atomic layer of ferromagnetic iron deposited on a non-magnetic substrate) closely follows the behaviour<sup>6</sup> of the two-dimensional Ising model<sup>7</sup>—the first and most elementary statistical model of a macroscopic system with short-range interactions<sup>3</sup>.

We grow epitaxial Fe films at room temperature on top of a non-magnetic single-crystal W(110) surface. Pressure in the vacuum chamber during growth does not exceed  $2 \times 10^{-10}$  mbar. In the ferromagnetically ordered phase, the films have a uniaxial easy magnetization axis along the in-plane [110] direction. Along this direction the magnetization *M* versus applied field *H* follows a square hysteresis loop, that is, within experimental uncertainty ( $< 1\%$ ) *M* at *H*=0 coincides with the value for *M* in an applied magnetic field. Accordingly<sup>8</sup>, such films are in a single domain state. *M* is measured by means of the magneto-optic Kerr effect<sup>9</sup>.

As revealed by our scanning tunnelling microscopy study of the epitaxial growth, films between 1 and 2 atomic layers (AL) consist of one complete AL of Fe followed by a two-dimensional

(2D) network of irregular patches, (Fig. 1). Some third layer is already visible. Low-energy electron diffraction shows, in this thickness range, a sharp  $p(1 \times 1)$  pattern indicating growth in registry with the substrate.

In the vicinity of the critical point  $[t = (T/T_c - 1), H] = [0, 0]$  the critical exponents  $\beta$  and  $\delta$  for the zero-field magnetization curve  $M(t, H=0) \propto (-t)^\beta$  and for the critical isotherm  $M(H, t=0) \propto H^{1/\delta}$  are determined to be  $0.13 \pm 0.02$  and  $14 \pm 5$  respectively, (Fig. 2). The theoretical 2D Ising values are  $\beta = 1/8$  (exactly known<sup>10</sup>) and  $\delta = 15$  (conjectured from numerical calculations<sup>11</sup>).

The sharpness of the transition observed in Fig. 2 is consistent with the magnetic correlation length reaching values of  $\sim 1,000$  nm (ref. 9). Evidently, the observed morphological defects—such as monatomic steps arising from the imperfect substrate (typical separation, 50–200 nm in the present sample) and the steps due to patches formation (Fig. 1) during growth—are not able to break the correlation length, that is, the correlations extend to much larger length than the spacing between defects. The picture emerging by comparing Fig. 1 and Fig. 2 is one where locally the number of coexisting layers (that is, the coordination number of Fe atoms) changes, thus causing local fluctuations of the strength of the various magnetic interactions.  $T_c$ , however, turns out to be sensitive to the average strength of the magnetic interactions over large distances rather than to the local fluctuations and thus is well defined. Notice that the lateral correlation length growing much larger than the film thickness is the key to observing a truly 2D phase transition.

Figure 3 shows a  $M$  versus  $t$  family of isochamps (curves with constant magnetic field). This figure shows the familiar picture  $M(t)$  curves separating out according to the strength of the applied magnetic field. In agreement with the high value of the exponent  $\delta$ , fields as small as 0.01 Oe (ref. 9) already introduce a measurable tail in the  $M(t)$  curve. This strong response to minute magnetic fields is characteristic of 2D systems, for which  $\delta = 15$  (ref. 11), and is not found in three-dimensional systems for which  $\delta = 5$  (ref. 5).

The family of isochamps given in Fig. 3 collapses into just one single curve—the scaling function (that is, the equation of

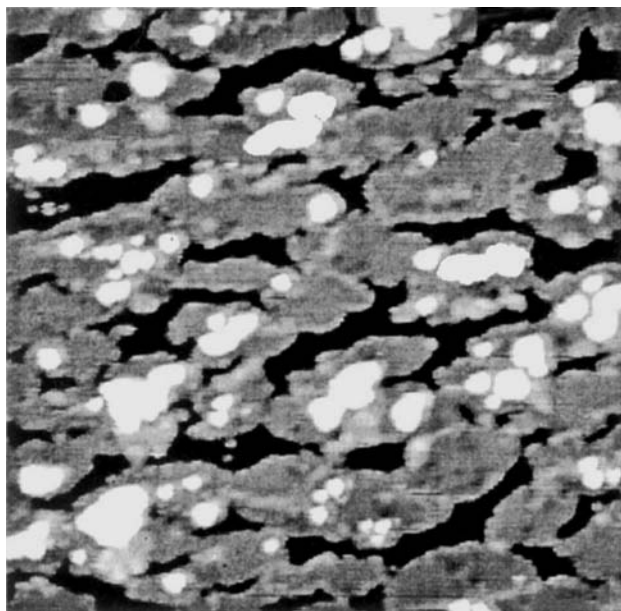


FIG. 1 STM topography of a  $1.8 \pm 0.1$  AL of Fe on W(110). The image shows an area of  $100 \times 100$  nm. The grey-scale range is 0.36 nm. Black corresponds to the first, grey to the second and white to the third Fe layer. A tunnelling current of 0.2 nA and a sample voltage of  $-620$  mV were used to collect the image.

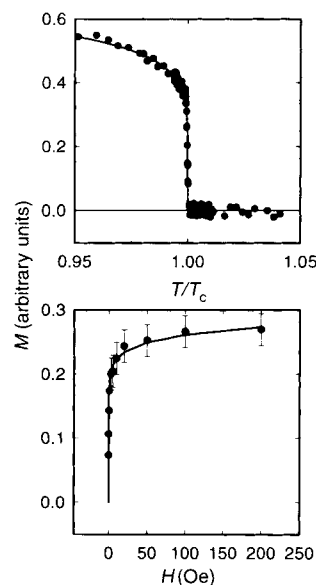


FIG. 2 Top, zero-field magnetization curve versus  $T/T_c$  ( $H = 0 \pm 10$  mOe) measured using the magneto-optical Kerr effect for an Fe film  $1.7 \pm 0.1$  AL thick. The thickness was determined from the STM topography. Absolute temperatures are determined to within  $\pm 5$  K, while relative temperatures are accurate to within 0.01 K.  $T_c$  (316.77 K in the present case) was determined separately, by locating the temperature at which  $[M(T, H) - M(T, 0)]$  has a maximum<sup>9</sup>, with a relative accuracy of  $\pm 0.07$  K. This means that the origin of the  $t$ -axis is known with an accuracy of  $\pm 2 \times 10^{-4}$ . The solid line is a power-law fit to the data  $M \propto (1 - T/T_c)^\beta$  yielding  $\beta = 0.13 \pm 0.02$ . The error in  $\beta$  is due to the uncertainty in locating  $T_c$ . Bottom, magnetization as a function of the applied magnetic field at  $T_c$ . The solid line is a power-law fit to the data  $M \propto H^{1/\delta}$ , yielding  $\delta = 14 \pm 5$ . The critical isotherm was obtained by plotting the maximum value of the family of curves  $[M(t, H) - M(t, 0)]$  versus  $H$ .

state in the critical region)—when the variables  $M$  and  $t$  are scaled by  $H^{1/\delta}$  and  $H^{1/\beta\delta}$ , respectively<sup>12</sup> (Fig. 4a). This collapsing is a manifestation of the scaling hypothesis<sup>1,12</sup>. Except for a discrete number of data points on the left-hand side of the figure, deviation from scaling is  $< 10\%$ .

Whereas the representation of Fig. 4a—introduced by Hankey and Stanley<sup>12</sup>—is most suited to visualize the data-collapsing, plotting<sup>13</sup>  $h(x) = H/M^\delta$  versus  $x = t/M^{1/\beta}$  allows the determination of the zero-field susceptibility critical exponents  $\gamma$  and of the ratio of the critical amplitudes  $\Gamma_+/\Gamma_-$ , defined as  $\chi(T) = \Gamma_+(t)^{-\gamma}$  ( $T > T_c$ ) and  $\chi(T) = \Gamma_-(-t)^{-\gamma}$  ( $T < T_c$ ), see Fig. 4b). This representation of scaling is designated as the  $h(x)$  representation of the scaling function<sup>13</sup>. For large  $x$  we expect the leading term<sup>14</sup> in  $h(x)$  to be  $x^\gamma/\Gamma_+$  (the point  $(h(\infty), \infty)$  corresponds to the critical isochore). In this representation, the point  $(h(x_0)=0, x_0)$  corresponds to the zero-field magnetization curve shown in Fig. 2. In the vicinity of this point one expects<sup>14</sup> the leading term in  $h(x)$  to be linear in  $(x+x_0)$  with coefficient  $\beta/\Gamma_-x_0^{(\gamma-1)}$ . The point  $(h(0), 0)$  is the critical isotherm shown in Fig. 2. By fitting the measured scaling function in the corresponding ranges according to the expected functional form we obtain  $\gamma = 1.74 \pm 0.05$  and  $\Gamma_+/\Gamma_- = 40 \pm 10$ . The theoretical 2D Ising values are  $7/4$  (exactly known<sup>15</sup>) and  $\Gamma_+/\Gamma_- \approx 37.7$  (ref. 16).

The coincidence between the measured critical quantities  $\beta$ ,  $\delta$ ,  $\gamma$ ,  $\Gamma_+/\Gamma_-$  and the calculated ones for the 2D Ising model might seem surprising at first sight. Our system is certainly not—see Fig. 1—the lattice with perfect translational symmetry underlying the Ising hamiltonian and the calculations by Gaunt and Domb<sup>6</sup>. Also Fe does not have localized spins as the Ising hamiltonian, which imitates rather an insulating ferromagnet with localized moments than a broad-band metallic ferromagnet like Fe. Moreover, although our uniaxial ferromagnet might

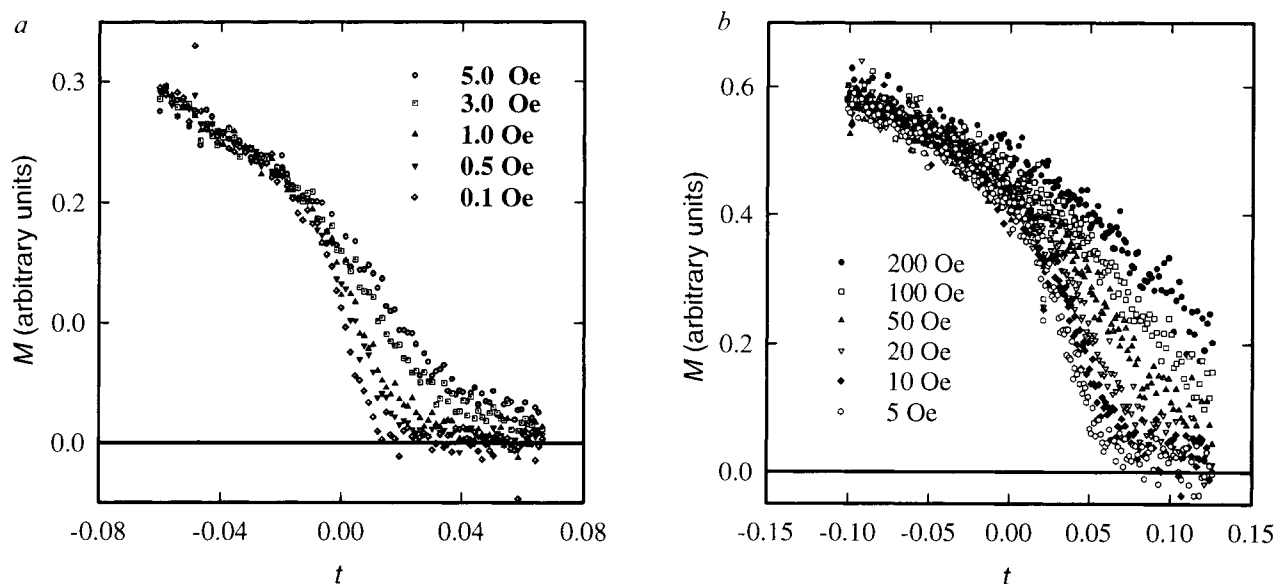


FIG. 3 Magnetization curves in constant magnetic field obtained in the vicinity of the Curie temperature as a function of  $t$ . For sample 1 (panel a) the applied field was varied between 0.1 and 5 Oe. For sample 2 (panel b) the field was varied between 5 and 200 Oe. The Curie temperatures were determined as in Fig. 2. The Curie temperature

of sample 1 was 331.3 K. Sample 2 was slightly thicker and had a Curie temperature of 345 K. Notice that in this thickness range ( $\sim 1.7$  AL) the phase transition occurs at temperatures close to deposition temperature, so the morphology does not change during the thermal scan.

have the same rotational symmetry as the 2D Ising hamiltonian, the strength of the various magnetic anisotropies in our films is very small (of the order of 1 K, compared to  $T_c \approx 300$  K). Accordingly, the individual Fe-spins can move easily away from the easy axis and thus have more degrees of freedom than the Ising spins, which can only be 'up' or 'down'. In fact, the temperature dependence of the magnetization far away from the critical point deviates strongly from the 2D Ising behaviour (ref. 9). Finally, a realistic magnetic interaction between two Fe

spins  $S_i S_j$  should be much more complicated than the scalar product ( $S_i, S_j$ ) forthcoming in the Ising hamiltonian. Thus, the good agreement between the experimental and measured critical quantities should be regarded as a stringent test of the universality hypothesis<sup>1,2</sup>. According to this hypothesis, it is believed<sup>14</sup> that the critical exponents and the ratio of critical amplitudes depend only on (1) the spatial dimensionality of the system, (2) the number of components of the order parameter, (3) the symmetry of the hamiltonian and (4) the range of the micro-

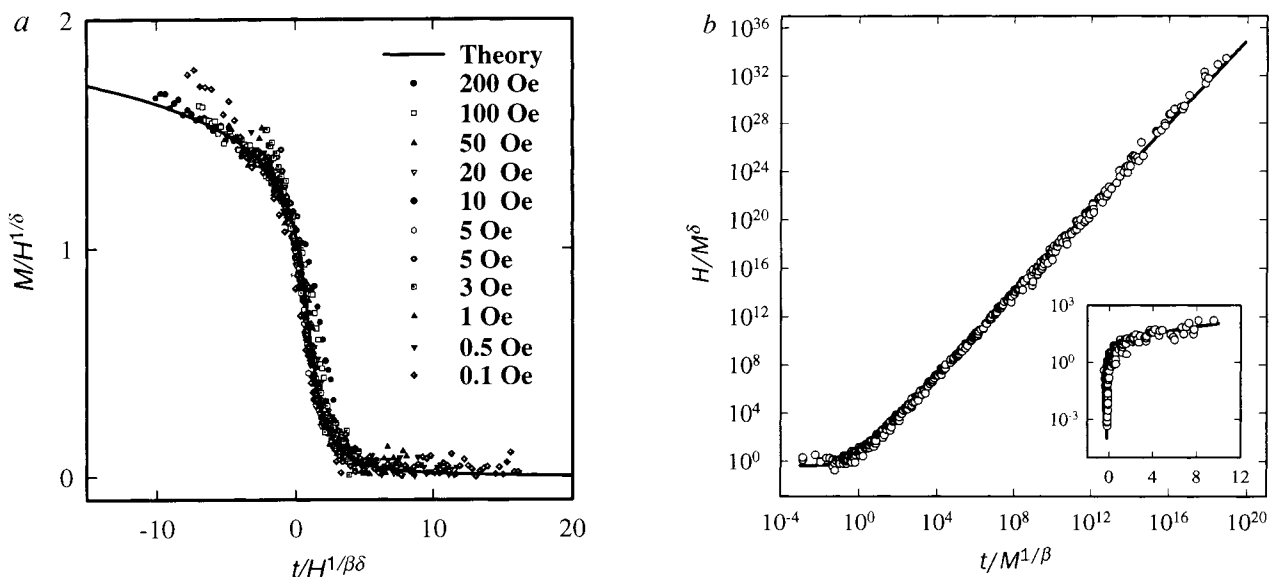


FIG. 4 a, Scaled magnetization  $M/H^{1/\delta}$  versus scaled temperature  $t/H^{1/\beta\delta}$  for the two films of Fig. 3. A linear scale is used. The solid line is the theoretical scaling function after ref. 6. For small fields, a discrete number of data points deviate from the scaling function. They are neglected in b. We used the value of  $\beta$  and  $\delta$  determined from Fig. 2. For  $\delta$  outside the range 12–16 the data collapsing is poorly realized.

This allows us to reduce the experimental uncertainty for  $\delta$  to  $\pm 2$ . b, Plot of  $H/M^\delta$  versus  $t/M^{1/\beta}$  for the two films of Fig. 3. The data with  $x > 0$  are represented as a log-log plot. The data points in the interval  $-1 < x < 10$  are represented semilogarithmically in the inset. The solid line is the theoretical curve from ref. 6. All data points, regardless of their value of  $H$ , are shown with the same symbol.



scopic interactions responsible for the phase transition. Also the presence of defects is expected to alter the universality class<sup>14</sup>. Apparently, the morphological defects observed in our system (Fig. 1) are not able to bring our system outside the Ising universality class. This might be related to the fact that such defects are not effective in limiting the correlation length.

A deeper level of universality was predicted<sup>1,3,5,14</sup>. At this level, provided that the non-universal scales for  $M$  and  $H$  are specified, the resulting scaling function should coincide within one universality class. This level of universality is also known historically as the Van der Waals Law of Corresponding States<sup>4</sup>: using properly normalized thermodynamic variables, all systems within the same universality class have the same equation of state. Thus, by suitably rescaling  $M$  and  $H$  (or, equivalently,  $h(x)$  and  $x$ ) it should be possible to bring the experimental scaling function and the theoretical one to coincidence. This rescaling is performed in Fig. 4, where the solid line is the theoretical scaling function of ref. 6 (the algebraic expressions used for constructing the solid lines in Fig. 4 are given explicitly in the Appendix of ref. 6). The theoretical and experimental results are normalized by the non-universal critical parameters  $h(0)$  and  $x_0$ , as suggested in ref. 12. The agreement between the experimental scaling function and the calculated one extends over 18 orders of magnitude of the variable  $x = t/M^{1/\beta}$ .

The exact solution of the 2D Ising model in an applied magnetic field is still missing<sup>2</sup>, so that one has to resort to approximate solutions. Our results confirm the validity of such calculations and represent a test of universality and scaling over

a wide range of parameters. Further experimental work on this system along the lines of for example, refs 17–21 (including the experimental determination of the scaling behaviour of the correlation function<sup>14</sup> by, for example, a new generation of neutron spectrometers) should improve our knowledge of cooperative phenomena beyond the pure numerical awareness. □

Received 21 July; accepted 31 October 1995.

1. Kadanoff, L. P. *Physica* **A163**, 1–14 (1990).
2. Baxter, R. J. *Exactly Solved Models in Statistical Mechanics* (Academic, London, 1982).
3. Levitt Sengers, A., Hocken, R. & Sengers, J. V. *Physics Today* **30**, 42–51 (1977).
4. De Boer, J. *Physica* **73**, 1–27 (1974).
5. Levitt Sengers, J. M. H. *Physica* **73**, 73–106 (1974).
6. Gaunt, D. S. & Domb, C. *J. Phys.* **C3**, 1442–1461 (1970).
7. Ising, E. *Z. Phys.* **31**, 253–258 (1925).
8. Kerkmann, D., Pescia, D. & Allenspach, R. *Phys. Rev. Lett.* **68**, 686–689 (1992).
9. Back, C. H., Würsch, C., Kerkmann, D. & Pescia, D. *Z. Phys.* **B96**, 1–3 (1994).
10. Yang, C. N. *Phys. Rev.* **85**, 808–816 (1952).
11. Domb, C. & Sykes, M. F. *Proc. R. Soc. Lond. A* **240**, 214–222 (1957).
12. Milosevic, S. & Stanley, H. E. in *Rendiconti S. I. F., Course LIX* (eds Müller, K. H. & Rigamonti, A.) (North-Holland, Amsterdam, 1976).
13. Griffiths, R. B. *Phys. Rev.* **158**, 176–87 (1967).
14. Kumar, A., Krishnamurthy, H. R. & Gopal, E. S. R. *Phys. Rep.* **98**, 57–143 (1983).
15. Abraham, D. B. *Phys. Lett.* **43A**, 163–164 (1973).
16. Barouch, E., McCoy, B. M. & Wu, T. T. *Phys. Rev. Lett.* **31**, 1409–1411 (1973).
17. Campuzano, J. C., Foster, M. S., Jennings, G., Willis, R. F. & Unertl, W. *Phys. Rev. Lett.* **54**, 2684–2687 (1985).
18. Qiu, Z. Q., Pearson, J. & Bader, S. D. *Phys. Rev. Lett.* **67**, 1646–1648 (1991).
19. Samulsen, E. J. *J. Phys. Chem. Solids* **35**, 785–793 (1974).
20. Als-Nielsen, J., Birgenau, R. J., Guggenheim, H. J. & Shirane, G. *Phys. Rev.* **B12**, 4963–4979 (1975).
21. Birgenau, R. J., Guggenheim, H. J. & Shirane, G. *Phys. Rev.* **B8**, 304–311 (1973).

ACKNOWLEDGEMENTS. This work was supported by the Schweizerischer Nationalfonds zur Förderung der wissenschaftliche Forschung.

## A 100-kyr periodicity in the flux of extraterrestrial <sup>3</sup>He to the sea floor

K. A. Farley & D. B. Patterson

Division of Geological and Planetary Sciences, MS 170-25, California Institute of Technology, Pasadena, California 91125, USA

MOST of the helium-3 in oceanic sediments comes from interplanetary dust particles (IDPs), and can therefore be used to infer the accretion rate of dust to the Earth through time<sup>1–3</sup>. <sup>3</sup>He records from slowly accumulating pelagic clays indicate that the accretion rate varies considerably over millions of years, probably owing to cometary and asteroidal break-up events<sup>3</sup>. Muller and MacDonald have proposed<sup>4</sup> that periodic changes in this accretion rate due to a previously unrecognized 100-kyr periodicity in the Earth's orbital inclination might account for the prominence of this frequency in climate records of the past million years<sup>5</sup>. Here we report variations in the <sup>3</sup>He flux to the sea floor that support this idea. We find that the flux recorded in rapidly accumulating Quaternary sediments from the Mid-Atlantic Ridge oscillates with a period of about 100 kyr. We cannot yet say, however, whether the 100-kyr climate cycle is a consequence of, a cause of, or an effect independent of these periodic changes in the rate of delivery of interplanetary dust to the sea floor.

The samples are from Deep Sea Drilling Project (DSDP) Site 607, on the flank of the Mid-Atlantic Ridge at 41° N. The core is a predominantly foraminiferal nanofossil ooze<sup>6</sup>, and carries a continuous, high-sedimentation-rate record between 245 kyr and 1.6 Myr (ref. 6). Variations in  $\delta^{18}\text{O}$ ,  $\delta^{13}\text{C}$ ,  $\text{CaCO}_3$  content and foraminifera populations in this core provide chronostratigraphic control and a detailed record of global climate conditions<sup>8</sup>. As with other climate records from deep-sea cores,

Site 607 demonstrates 23/19- and 41-kyr climate periodicities and the onset of 100-kyr glacial-interglacial cycling about 900 kyr ago. The 23/19- and 41-kyr periodicities can be attributed to precession and obliquity of the Earth's orbit, but the origin of the 100-kyr signal remains controversial<sup>5</sup>. Although variations in orbital eccentricity are of roughly the appropriate frequency<sup>7</sup>, the resulting insolation modulation is insufficient to account for the observed climate record<sup>5</sup>, nor is it clear why this signal first appeared about a million years ago. Identification of the origin of this cycle remains central to understanding Earth's climate system<sup>7</sup>. We have analysed the <sup>3</sup>He content of a complete 9-m section of core spanning 253 to 458 kyr to characterize the accretion rate of extraterrestrial matter over 10<sup>5</sup>-year timescales and to investigate its possible relationship<sup>4</sup> to this 100-kyr climate periodicity.

Unlike previously studied sediments<sup>1–3,9</sup>, this core is characterized by fairly low <sup>3</sup>He/<sup>4</sup>He ratios of  $\sim 2 \times 10^{-7}$  (Table 1), reflecting accumulation of <sup>4</sup>He-rich terrigenous material. Nevertheless, the <sup>3</sup>He/<sup>4</sup>He ratios are nearly an order of magnitude higher than typical crustal matter, demonstrating the presence of an additional <sup>3</sup>He-rich component. To identify this <sup>3</sup>He source, we performed additional experiments on the samples. Magnetic separates were found to be highly enriched in <sup>3</sup>He and <sup>3</sup>He/<sup>4</sup>He ratio (by up to 30 times over the bulk sediment), and stepped heating experiments further confirmed that <sup>3</sup>He and <sup>4</sup>He reside in different carriers. <sup>4</sup>He is overwhelmingly released at low temperatures (<500 °C), whereas <sup>3</sup>He has a bimodal release, at <500 °C and 950 °C. The high-temperature release (>600 °C) accounts for ~75% of the total <sup>3</sup>He. Helium release peaking at 950 °C and enrichment of <sup>3</sup>He in magnetic fines are the hallmark of IDP helium within sediments<sup>10,11</sup>. Using a simple deconvolution model that attributes the <sup>3</sup>He to crustal, atmospheric and extraterrestrial components, we calculate that in all samples, more than 75% of the <sup>3</sup>He is extraterrestrial (range 75–92%; Table 1). Replicate <sup>3</sup>He measurements have a typical standard deviation of 15%, far larger than our analytical uncertainty of ~2%. This variability is consistent with the presence of a rela-

tively small number of helium-rich IDPs within the replicates. By analysing multiple 0.5-g aliquots, the standard deviation of the mean of each analysed core segment has been brought down to about 10%.

The concentration of extraterrestrial  $^3\text{He}$  ( $^3\text{He}_{\text{ET}}$ ) fluctuates smoothly down the core over a total range of about a factor of three (Fig. 1a). This quantity is controlled by the flux of extraterrestrial  $^3\text{He}$  to the sea floor ( $f$ ), the  $^3\text{He}$  retentivity ( $R$ ), and the sediment-mass accumulation rate ( $\alpha$ ) by the relation  $^3\text{He}_{\text{ET}} = fR/\alpha$  (ref. 3). By tuning the  $\delta^{18}\text{O}$  record to the obliquity and precessional rhythms, Ruddiman *et al.*<sup>8</sup> identified eleven  $\sim 23$ -kyr precessional cycles in the length of core we studied, and these constrain the sedimentation rate. Our samples represent a continuous and thoroughly homogenized sediment strip from the entire length of each precessional cycle<sup>7</sup>, so  $\alpha$  can be obtained from the linear sedimentation rate and the mean dry bulk density of the analysed segment (Table 1). Because extraterrestrial  $^3\text{He}$  is strongly retained in sediments for  $>70$  Myr (ref. 3), variations in retentivity are unlikely to be important over the short period recorded here.

The average flux of extraterrestrial  $^3\text{He}$  to the sea floor during each precessional cycle is shown in Fig. 1b. These are truly representative average fluxes because we have sampled a strip from each entire cycle, so short-term events are not over- or under-represented as they might be by spot sampling. The mean flux over the eleven cycles of  $1.07 \times 10^{-12} \text{ cm}^3 \text{ STP cm}^{-2} \text{ kyr}^{-1}$  is indistinguishable from previous estimates for the mean Quaternary flux<sup>2,3,9</sup>, and substantiates our deconvolution technique. The flux varies significantly and smoothly down-core, with maxima near 335 and possibly 425 kyr, and minima at 280 and 385 kyr. The total variability is 350%, far larger than the  $\sim 10\%$  uncertainty on  $^3\text{He}_{\text{ET}}$  in each cycle. Although the record is short, Fig. 1b strongly suggests that the flux oscillates with a period of  $\sim 100$  kyr, similar to both orbital variations and to climate.

Figure 1 compares the  $^3\text{He}$  flux to orbital parameters and to global climate as recorded by foraminiferal  $\delta^{18}\text{O}$  at Site 607.  $^3\text{He}$  fluxes are positively correlated with inclination relative to the invariable plane<sup>4</sup> and, to a lesser extent, with eccentricity<sup>7</sup>. In addition, the  $^3\text{He}$  flux covaries with the 100-kyr component of the climate signal. Peak interglacial conditions (low  $\delta^{18}\text{O}$ ; maxima in Fig. 1c) at both 325 and 410 kyr are preceded by at least 35 kyr of high  $^3\text{He}$  flux. The termination of interglacial peaks is coincident with the decline in  $^3\text{He}$  flux. This similarity raises the possibility that the calculated flux variation is an artefact of an improper compensation of sedimentation rate effects—that is, the variations in flux may be a spurious product of errors in core chronostratigraphy that are somehow correlated with climate. In order to make the flux variations disappear, the sedimentation rate would have to vary inversely with the observed  $^3\text{He}$  concentration, rather than directly as the chronostratigraphic model suggests (Table 1). Relative errors in the assigned ages would have to be up to 17 kyr and would have to vary in a complex manner down the core. Such errors seem implausibly large, given the quality of the Site 607  $\delta^{18}\text{O}$  fit to other dated sites<sup>12</sup> and to the Milankovitch 23/19- and 41-kyr forcings<sup>8</sup>. On similarly tuned cores, the estimated error is better than 5 kyr (ref. 13).

A very important question is what, if any, causal relationships exist between the 100-kyr oscillations observed in orbital parameters, climate and  $^3\text{He}$  flux. We consider three possibilities: (1) the 100-kyr climate cycle controls the delivery of IDPs to the sea floor; (2) orbital parameters modulate the IDP accretion rate, which in turn influences climate<sup>4</sup>; and (3) the IDP accretion rate and global climate are correlated but otherwise independent responses to orbital inclination and/or eccentricity.

The first possibility requires no causal relationship between orbital parameters and the observed  $^3\text{He}$  flux. Instead, 100-kyr climate effects could control the observed  $^3\text{He}$  flux by focusing of IDPs to particular regions of the sea floor by oceanic currents, ice rafting, or wind. IDPs that retain their helium are likely to be a few tens of micrometres in diameter<sup>1</sup>. The Stokes settling

time of such particles through the oceanic water column is less than a decade, so focusing during descent to the sea floor is unlikely to be significant, as is riverine transport from the continents. Resuspension and focused redeposition of extraterrestrial particles by bottom currents is a more likely source of local variations in  $^3\text{He}$  flux, and although preservation of climate proxy stratigraphy argues against large-scale remobilization at Site 607, some fine-scale transport is possible<sup>6</sup>. Ice rafting might be expected to lead to higher  $^3\text{He}$  deposition during glacial periods, but the opposite effect is observed in our record. In the atmosphere, the residence time for particles in the appropriate size range is probably a few days to weeks<sup>14</sup> and particulates of  $\sim 20 \mu\text{m}$  diameter injected into the stratosphere have been observed to travel up to a few thousand kilometres<sup>15</sup>. If removal from the stratosphere is highly localized, then some wind-focusing may be possible. Because the intensity of these mechanisms may be related to climate conditions, it is possible that glacial-interglacial cycling may modulate the distribution of IDPs on the Earth's surface. The variations in Fig. 1b would require very efficient and climate-sensitive geographic focusing.

In contrast, the second and third possibilities require a causal link between the observed  $^3\text{He}$  flux and orbital parameters, through modulation of the accretion rate of interplanetary dust. It is known from satellite observations that interplanetary dust is inhomogeneously distributed in space, being localized in bands

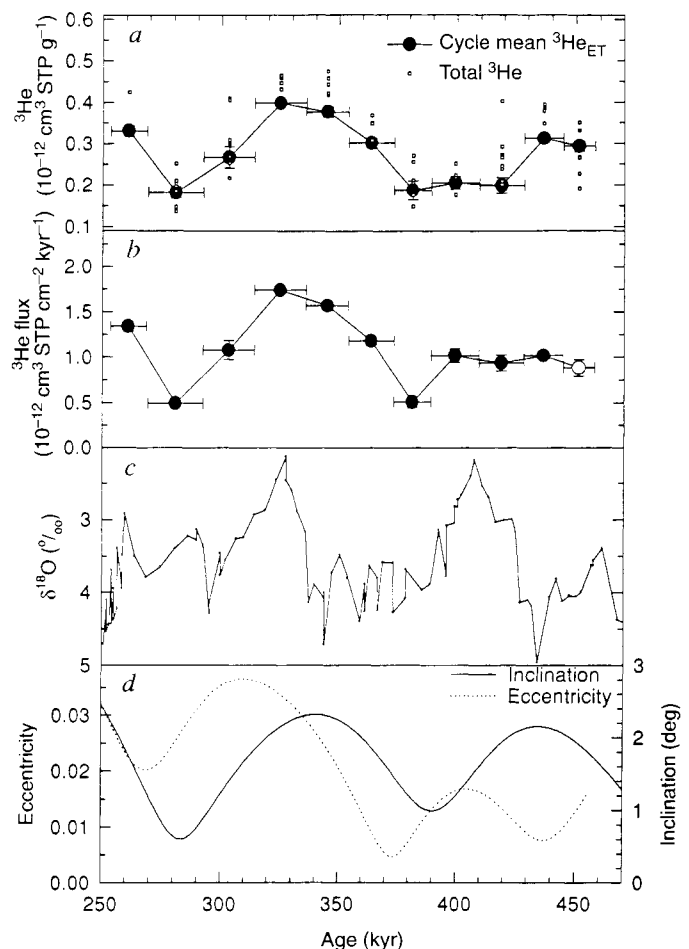


FIG. 1 a, Measured  $^3\text{He}$  concentrations and precession-cycle average  $^3\text{He}_{\text{ET}}$  in Hole 607 sediments; b, extraterrestrial  $^3\text{He}$  flux computed from the cycle-average data in a and  $\alpha$ -listed in Table 1 (cycle 10 is indicated by an open symbol to highlight the large uncertainty associated with its  $\alpha$  value). c, Oxygen isotope climate proxy, from ref. 8; d, orbital eccentricity<sup>7</sup> and inclination relative to the invariable plane<sup>4</sup>.

TABLE 1 DSDP Hole 607 helium results

| Cycle | Age<br>(kyr) | <i>n</i> | $^3\text{He}$<br>( $10^{12} \text{ cm}^3$<br>STP $\text{g}^{-2}$ ) | $\pm$ | $^3\text{He}/^4\text{He}$<br>( $\times 10^{-6}$ ) | $\pm$ | ET<br>% | $^3\text{He}$ flux<br>( $10^{12} \text{ cm}^3$<br>STP $\text{cm}^{-2} \text{ kyr}^{-1}$ ) | $\pm$ | $\alpha$<br>( $\text{g cm}^{-2} \text{ kyr}^{-1}$ ) |
|-------|--------------|----------|--------------------------------------------------------------------|-------|---------------------------------------------------|-------|---------|-------------------------------------------------------------------------------------------|-------|-----------------------------------------------------|
| 0     | 253–268      | 1        | 0.424                                                              | 0.010 | 0.091                                             | 0.005 | 78      | 1.34                                                                                      | 0.05  | 4.05                                                |
| 1     | 268–292      | 4        | 0.208                                                              | 0.014 | 0.173                                             | 0.013 | 87      | 0.49                                                                                      | 0.04  | 2.76                                                |
| 2     | 292–314      | 7        | 0.306                                                              | 0.028 | 0.161                                             | 0.015 | 87      | 1.08                                                                                      | 0.11  | 4.06                                                |
| 3     | 314–336      | 4        | 0.443                                                              | 0.009 | 0.204                                             | 0.004 | 89      | 1.74                                                                                      | 0.04  | 4.31                                                |
| 4     | 336–354      | 4        | 0.440                                                              | 0.014 | 0.127                                             | 0.005 | 85      | 1.57                                                                                      | 0.07  | 4.70                                                |
| 5     | 354–373      | 4        | 0.346                                                              | 0.014 | 0.157                                             | 0.006 | 86      | 1.18                                                                                      | 0.05  | 3.62                                                |
| 6     | 373–389      | 5        | 0.211                                                              | 0.023 | 0.195                                             | 0.018 | 89      | 0.51                                                                                      | 0.06  | 2.72                                                |
| 7     | 389–409      | 4        | 0.223                                                              | 0.018 | 0.281                                             | 0.018 | 92      | 1.02                                                                                      | 0.08  | 4.85                                                |
| 8     | 409–428      | 8        | 0.268                                                              | 0.022 | 0.089                                             | 0.001 | 75      | 0.94                                                                                      | 0.10  | 4.66                                                |
| 9     | 428–445      | 4        | 0.374                                                              | 0.010 | 0.124                                             | 0.005 | 82      | 1.02                                                                                      | 0.03  | 3.31                                                |
| 10    | 445–458      | 6        | 0.273                                                              | 0.023 | 0.104                                             | 0.007 | 80      | 0.88                                                                                      | 0.09  | 4.01*                                               |

Site 607 is located at  $41^\circ \text{N}$ ,  $33^\circ \text{W}$ , 3,420 m depth<sup>6</sup>. Relative precessional cycle number is indicated (cycles 0 and 10 are incomplete owing to coring problems). Ages are from ref. 8. Analysed samples are a complete and homogenized composite of each indicated age interval. *n*, Number of  $\sim 500$  mg replicates (except cycle 0, from which a single 1.8-g sample was analysed), uncertainty figures are  $1\sigma$  of the mean from the *n* analyses (analytical uncertainties alone are about 2% for both  $^3\text{He}$  and  $^4\text{He}$ ). Uncertainties in the calculated flux only reflect the uncertainties in  $^3\text{He}$  averages. Results are corrected for blanks of  $0.6 \times 10^{-15} \text{ cm}^3$  STP of  $^3\text{He}$  and an atmospheric  $^3\text{He}/^4\text{He}$  ratio.  $\alpha$ , Sediment-mass accumulation rate calculated from the age ranges and sediment density (from ref. 6; supplemented with densities from  $\text{CaCO}_3$  abundance<sup>8</sup> following the procedure described in ref. 18 as necessary). ET %, fraction of  $^3\text{He}$  that is extraterrestrial, calculated as follows:  $^3\text{He}_m = ^3\text{He}_{\text{air}} + ^3\text{He}_c + ^3\text{He}_{\text{ET}}$ ,  $^4\text{He}_m = ^4\text{He}_{\text{air}} + ^4\text{He}_c + ^4\text{He}_{\text{ET}}$ , where the subscripts m, air, c, and ET refer to measured, air-derived, crustal and extraterrestrial respectively. In these sediments extraterrestrial  $^4\text{He}$  can be ignored. The abundance of air-derived helium can be calculated and corrected for by measuring neon abundances in the sample and by assuming addition of He and Ne in atmospheric proportions. The measured Ne in these samples is extraordinarily low, and no correction for atmospheric helium is actually necessary. The decay of U and Th as well as cosmic-ray spallation produce both  $^4\text{He}$  and  $^3\text{He}$  in crustal rocks. The average crustal  $^3\text{He}/^4\text{He}$  ratio ( $^3\text{He}/^4\text{He}_c$ ) is reasonably well known ( $\sim 2 \times 10^{-6}$ ; refs 19, 20 and our unpublished measurements on loess samples) and  $^3\text{He}_c = (^3\text{He}_m - ^3\text{He}_{\text{air}}) \times (^3\text{He}/^4\text{He}_c)$ . Although some crustal rocks with higher  $^3\text{He}/^4\text{He}$  ratios exist as a consequence of long cosmic-ray exposure histories, we do not believe this source contributes more  $^3\text{He}$  than implied by the mean crustal  $^3\text{He}/^4\text{He}$  ratio. This conclusion is supported by our observation of high-temperature release of sediment  $^3\text{He}$ , rather than low-temperature release characteristic of cosmogenic  $^3\text{He}$  (ref. 21). Furthermore, the mean exposure ages required to produce the  $^3\text{He}$  in the non-carbonate fraction of these samples are  $>1$  Myr, far higher than is geologically reasonable. Note that selection of a different but reasonable crustal value would make little difference to our conclusions. DSDP hole 607 sample designations for each cycle are as follows: (0) 2-1 0-91; (1) 2-1, 88–150 cm + 2-2, 0–8 cm; (2) 2-2, 8–106 cm; (3) 2-2, 106–150 cm + 2-3, 0–53 cm; (4) 2-3, 53–150 cm; (5) 2-4, 0–78 cm; (6) 2-4, 78–121 cm; (7) 2-4, 121–150 cm + 2-5, 0–69 cm; (8) 2-5, 69–150 cm + 2-6, 0–8 cm; (9) 2-6, 8–91 cm; (10) 2-6, 91–150 cm + 2-7, 0–26 cm.

\* Core disruption between cores 607-2 and 607-3 precludes accurate determination of sedimentation rate in this interval; the mean value of the other 10 cycles has been adopted for this cycle.

related to asteroid families<sup>16</sup>, and confined in rings gravitationally locked to the Earth's orbit<sup>17</sup>. Without attempting to specify a mechanism, we observe that it is plausible that variations in the Earth's orbit induce changes in the rate at which extraterrestrial dust is accreted. Our data would suggest higher accretion rates during periods of high inclination and/or eccentricity.

Discrimination between possibility (1), and possibilities (2) and (3) can be made by analysis of other sediment cores. Variations in the extraterrestrial accretion rate should leave a similar record, in phase and magnitude, in cores from around the world. In contrast, geographical focusing would be expected to leave much more variable records. There is only one other data set with which this test can be made. Marcantonio *et al.*<sup>9</sup> studied the most recent 200 kyr of a core from the eastern equatorial Pacific. Although their methodology and objective were different from ours, these authors detected two 10–20-kyr episodes of anomalously low  $^3\text{He}$  flux, at about 100 and 190 kyr. The Pacific record does not overlap in time with the Site 607 data, but it suggests that the lowest fluxes in these two cores occur at about 100, 190, 280 and 385 kyr, so with roughly a 100-kyr recurrence interval. These data suggest in-phase variations in different geographical and climate environments, and therefore support orbital modulation of the IDP accretion rate.

The available data are consistent with climate forcing by the accretion of extraterrestrial dust, although they by no means prove such causality. However, there is one additional and intriguing piece of evidence that may bear on this point. Previous work<sup>3</sup> has suggested that the IDP accretion rate rose dramatically between 3 and 1.1 Myr, probably as a consequence of increased cometary activity or asteroid collisions. This rise is essentially coincident with the onset of the 100-kyr glacial-interglacial periodicity at  $\sim 0.9$ –1 Myr (refs 5, 8). Because orbital parameters have varied in substantially the same manner over

the last few Myr (ref. 7), the coincident onset of both a higher IDP flux and 100 kyr climate periodicity cannot be attributed to orbital effects. Although a completely fortuitous synchronicity of these two uncommon events cannot be ruled out, the alternative explanation is that the accretion of extraterrestrial matter to the Earth (or its abundance in near-Earth space) rose dramatically  $\sim 1$  Myr ago and now drives the 100-kyr glacial-interglacial cycle. Although possible mechanisms for such a linkage have been suggested<sup>4</sup>, this issue remains problematic. Our data require that any mechanism must predict a direct correlation with eccentricity and/or inclination and must operate with an accretion rate variation of about a factor of 3.

Our preferred explanation for the  $^3\text{He}$  record at Site 607 is that it reflects variations in the accretion rate of interplanetary dust to the Earth, probably in response to the 100-kyr inclination and/or eccentricity cycles. Although speculative, the correlation of this periodicity with the 100-kyr glacial-interglacial cycle, together with the synchronous onset of the 100-kyr glacial periodicity and the sharp rise in the  $^3\text{He}$  flux about 1 Myr ago<sup>3</sup>, suggest that extraterrestrial dust may be involved in global climate. Whatever its origin, the existence of 100-kyr oscillations in the flux of  $^3\text{He}$  to the sea floor may provide additional insights to global climate processes.  $\square$

Received 19 June; accepted 30 October 1995.

- Ozima, M., Takayanagi, M., Zashu, S. & Amari, S. *Nature* **311**, 449–451 (1984).
- Takayanagi, M. & Ozima, M. *J. geophys. Res.* **92**, 12531–12538 (1987).
- Farley, K. A. *Nature* **376**, 153–156 (1995).
- Muller, R. A. & MacDonald, G. J. *Nature* **377**, 107–108 (1995).
- Imbrie, J. *et al.* *Paleoceanography* **8**, 699–735 (1993).
- Ruddiman, W. F. *et al.* *Init. Rep. DSDP Leg 94* (1987).
- Berger, A. & Loutre, M. F. *Quat. Sci. Rev.* **10**, 297–317 (1991).
- Ruddiman, W. F., Raymo, M. E., Martinson, D. G., Clement, B. M. & Backman, J. *Paleoceanography* **4**, 353–412 (1989).



9. Marcantonio, F., Kumar, N., Stute, M., Anderson, R. F. & Schlosser, P. *Earth planet. Sci. Lett.* **133**, 549–555 (1995).
10. Amari, S. & Ozima, M. *Nature* **317**, 520–522 (1985).
11. Amari, S. & Ozima, M. *Geochim. cosmochim. Acta* **52**, 1087–1095 (1988).
12. Shackleton, N. J., Berger, A. & Peltier, W. R. *Trans. R. Soc. Edinb. Earth Sci.* **81**, 251–261 (1990).
13. Martinson, D. G. et al. *Quat. Res.* **27**, 1–29 (1987).
14. Toon, O. B. et al. *Geol. Soc. Am. Spec. Pap.* **190**, 187–200 (1982).
15. Sigurdsson, H. *Geol. Soc. Am. Spec. Pap.* **247**, 99–110 (1990).
16. Sykes, M. V. et al. in *Asteroids II* (eds Binzel, R. P., Gehrels, T. & Matthews, M. S.) 336–367 (University of Arizona Press, Tucson, 1989).
17. Dermott, S. F., Jayaraman, S., Xu, Y., Gustafson, B. A. S. & Liou, J. C. *Nature* **369**, 719–723 (1994).
18. Herbert, T. D. & Mayer, L. A. *J. sedim. Petrol.* **61**, 1089–1108 (1991).
19. Tolstikhin, I. N. in *Terrestrial Rare Gases* (eds Alexander, E. C. & Ozima, M.) 33–62 (Japan Scientific Societies, Tokyo, 1978).
20. Andrews, J. N. *Chem. Geol.* **49**, 339–351 (1985).
21. Trull, T. W., Kurz, M. D. & Jenkins, W. J. *Earth planet. Sci. Lett.* **103**, 241–256 (1991).

ACKNOWLEDGEMENTS. This work was supported by the NSF. We thank W. Alvarez for early discussions which led to this work, R. Muller, F. Marcantonio, A. Anbar and S. Love for helpful discussions, R. Wolf for analytical assistance, and S. Dermott, R. Poreda and W. Ruddiman for careful reviews.

## Atmospheric carbon dioxide concentrations before 2.2 billion years ago

Rob Rye, Phillip H. Kuo & Heinrich D. Holland

Department of Earth and Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge, Massachusetts 02138, USA

THE composition of the Earth's early atmosphere is a subject of continuing debate<sup>1–4</sup>. In particular, it has been suggested that elevated concentrations of atmospheric carbon dioxide would have been necessary to maintain normal surface temperatures in the face of lower solar luminosity in early Earth history<sup>5,6</sup>. Fossil weathering profiles, known as palaeosols, have provided semi-quantitative constraints on atmospheric oxygen partial pressure ( $p_{O_2}$ ) before 2.2 Gyr ago<sup>36,37</sup>. Here we use the same well studied palaeosols to constrain atmospheric  $p_{CO_2}$  between 2.75 and 2.2 Gyr ago. The observation that iron lost from the tops of these profiles was reprecipitated lower down as iron silicate minerals<sup>7–9</sup>, rather than as iron carbonate, indicates that atmospheric  $p_{CO_2}$  must have been less than  $10^{-1.4}$  atm—about 100 times today's level of 360 p.p.m., and at least five times lower than that required in one-dimensional climate models to compensate for lower solar luminosity at 2.75 Gyr. Our results suggest that either the Earth's early climate was much more sensitive to increases in  $p_{CO_2}$  than has been thought, or that one or more greenhouse gases other than  $CO_2$  contributed significantly to the atmosphere's radiative balance during the late Archaean and early Proterozoic eons.

The mobility of iron during the weathering of basalts before 2.2 Gyr indicates that  $p_{O_2}$  was lower than  $\sim 2 \times 10^{-3}$  atm (ref. 9). During the anoxic weathering of basalts,  $Fe^{2+}$ , together with  $K^+$ ,  $Na^+$ ,  $Mg^{2+}$ ,  $Ca^{2+}$ ,  $H_4SiO_4$ , and  $HCO_3^-$ , is washed down into the lower parts of the weathering profile. The  $Fe^{2+}$  is then either lost to ground water or is precipitated as a constituent of one or more minerals. If an  $Fe^{2+}$ -bearing mineral precipitates, it will be siderite ( $FeCO_3$ ) if  $p_{CO_2}$  is high, and one or more iron silicates if  $p_{CO_2}$  is low. None of the four pre-2.2-Gyr palaeosols studied to date<sup>7,9,12,38</sup> contains siderite. Iron-rich chlorite is abundant in their lower sections, indicating that all or part of the  $Fe^{2+}$  released during weathering from the upper part of the palaeosols was reprecipitated in these soils during downward passage. The iron-rich chlorites are almost certainly products of diagenesis and/or low-grade metamorphism. Their precursors were probably  $Fe^{2+}$ -silicate and/or  $Fe^{2+}$ -aluminium silicate phases. It is unlikely that the chlorites were formed by the reaction of siderite with an aluminium silicate phase. Analyses of recent siderite samples and their associated pore waters indicate that if siderite had precipitated in the lower parts of these palaeosols from

solutions with an  $FeO/(FeO + MgO)$  ratio approximately equal to that of the parent basalts, the siderites would have been almost pure  $FeCO_3$  (refs 10, 11). Reaction of this siderite with an aluminosilicate mineral would have produced a nearly pure iron chlorite. Chlorites in the 2.2-Gyr Hekpoort palaeosol, which developed on basalt, and in a  $\sim 2.5$ -Gyr palaeosol in Canada, which developed on a mafic dyke<sup>12</sup>, have  $FeO^*/(FeO^* + MgO)$  ratios between 0.49 and 0.84 (Fig. 1). Chlorites of this composition are more likely to be products of the diagenesis of Fe–Mg clay minerals than of siderite. This interpretation is also favoured by the persistence of siderite in rocks metamorphosed at pressures and temperatures considerably higher than those experienced by the above palaeosols<sup>13</sup>.

The precipitation of  $Fe^{2+}$  as a constituent of silicates rather than carbonates sets an upper limit on the  $CO_2$  partial pressure in the soil solutions. At equilibrium between siderite and greenalite, the simplest of the possible iron silicates,  $Fe_3Si_2O_5(OH)_4 + 3CO_2 + 2H_2O \leftrightarrow 3FeCO_3 + 2H_4SiO_4$ , and the equilibrium constant  $K_{eq} = (a_{H_4SiO_4})^2 / (p_{CO_2})^3$ , where  $a$  is activity. The thermodynamic data of refs 14–16 can be used to calculate the position of the boundary between the stability fields of greenalite and siderite (Fig. 2). As Eugster and Chou's data<sup>14</sup> for greenalite are rather insecure, we have sought to confirm the lower limit of the stability of siderite in siliceous sediments by measuring the  $p_{CO_2}$  in ground waters from sideritic iron formations and by calculating the  $p_{CO_2}$  of ground waters in siderite-bearing sandstones and shales<sup>17–23</sup>. The results (Fig. 2) show that at temperatures of  $\sim 10^\circ C$  siderite is stable in the presence of iron silicate and aluminosilicate minerals at  $CO_2$  pressures down to  $\sim 10^{-2.2}$  atm. Although the data in Fig. 2 are not definitive, they suggest that  $p_{CO_2}$  in ground waters at  $10^\circ C$  in the lower parts of  $\geq 2.2$ -Gyr palaeosols was  $\leq 10^{-2.2}$  atm.

In modern soils,  $p_{CO_2}$  is nearly always higher than atmospheric  $p_{CO_2}$ , because the decay of organic matter adds  $CO_2$  to soil air. In the absence of land plants, soil  $p_{CO_2}$  decreases as water penetrates into weathering profiles. Soil  $p_{CO_2}$  is determined by the rate of diffusion through the soil and by the rate of reaction of  $CO_2$  with the soil. If  $CO_2$  diffusion is rapid, the downward

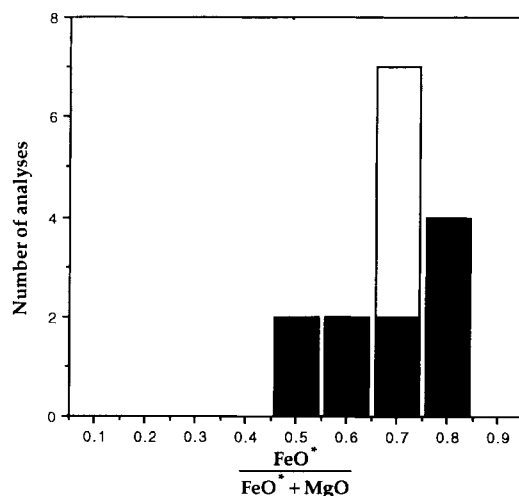


FIG. 1 Histogram of  $FeO^*/(FeO^* + MgO)$  ratios in palaeosol chlorites. Black bar: unpublished electron microprobe analyses by E. A. Zbinden and T. J. Hart of 10 chlorite samples from the 2.2 Gyr Hekpoort palaeosol, Transvaal, South Africa. White bar: published electron microprobe analyses of 5 chlorite samples from a  $\sim 2.5$ -Gyr palaeoweathering profile located north of Thessalon, Ontario, Canada, and stratigraphically below the Matinenda Formation<sup>12</sup>.  $FeO^* = FeO + Fe_2O_3$ .

decrease of  $p_{\text{CO}_2}$  is small. If  $\text{CO}_2$  diffusion is slow,  $p_{\text{CO}_2}$  can decrease significantly toward the base of soil profiles.

If the temperature of ground water in these palaeosols was  $>10^\circ\text{C}$ , the maximum value of  $p_{\text{CO}_2}$  in siderite-free palaeosols could have been  $>10^{-2.2}$  atm. Eugster and Chou's<sup>14</sup> thermodynamic data suggest that the maximum  $\text{CO}_2$  pressure in these palaeosols would rise by about 0.12 log units per  $5^\circ\text{C}$ . This is consistent with the slope determined by May for  $p_{\text{CO}_2}$  at the boundary between siderite- and chlorite-bearing assemblages in geothermal waters of the Rhenish Massif<sup>24</sup>.

The mean annual temperature in the world's warmest location today is  $31^\circ\text{C}$ . At this extreme temperature, siderite-free palaeosols could have developed in ground waters with  $p_{\text{CO}_2} \leq 10^{-1.7}$  atm. Pinto and Holland<sup>25</sup> have shown that in soils with a  $\text{CO}_2$  diffusion coefficient of  $2 \times 10^{-2} \text{ cm}^2 \text{ s}^{-1}$  and with typical reaction rates between soil water and unweathered parent material,  $p_{\text{CO}_2}$  can be expected to fall to about 60% of its surface value at a depth of 10 m. In the unlikely event that there is absolutely no downward diffusion of  $\text{CO}_2$  within the soil, the  $\text{CO}_2$  pressure in soil water decreases by a factor of about three before the soil water reaches saturation with respect to greenalite, when atmospheric  $p_{\text{CO}_2}$  is between  $10^{-1.0}$  and  $10^{-1.5}$  atm. At higher atmospheric  $\text{CO}_2$  levels, the fractional loss of  $\text{CO}_2$  during the downward passage of water before greenalite saturation is smaller.

We therefore conclude that the most likely upper limit to atmospheric  $p_{\text{CO}_2}$  in siderite-free palaeosols developed on basaltic rocks before 2.2 Gyr is  $10^{-1.4}$  atm (Fig. 3). This  $\text{CO}_2$  pressure at 2.75 Gyr, the age of the well studied Mt Roe palaeosols in Western Australia<sup>9</sup>, is inconsistent with cloud-free, one-dimensional climate models in which the greenhouse effect of

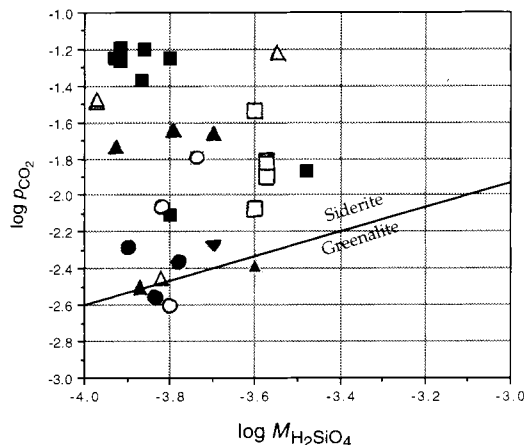


FIG. 2 Plot of  $\log p_{\text{CO}_2}$  (in atmospheres) versus  $\log M_{\text{H}_4\text{SiO}_4}$  (where  $M_{\text{H}_4\text{SiO}_4}$  is the molarity of dissolved silica) in ground waters from siderite-bearing formations and the calculated boundary between the stability fields of siderite and greenalite based on Eugster and Chou's<sup>14</sup> estimate of the Gibbs free energy of formation of greenalite ( $-710 \text{ kcal mol}^{-1}$ ) and a value of  $10^{-10.80}$  for the  $K_{\text{sp}}$  of siderite. This  $K_{\text{sp}}$  is consistent with the measurements reported in refs 15, 16. Filled circles, our analyses of samples from a number of municipal wells that draw from the Biwabik Formation in the Mesabi Iron Range of Northeastern Minnesota; open circles, literature data from the Biwabik<sup>17</sup>; filled squares, the Breathitt Formation in Eastern Kentucky<sup>18,19</sup>; open squares, the Norton Formation, Western Virginia<sup>20</sup>; filled inverted triangle, the Connoquenessing Formation, Pottsville Group<sup>21,22</sup>; open triangles, the Kittanning Formation, Allegheny Group<sup>21-23</sup>; filled triangles, the Freeport Formation, Allegheny Group<sup>23</sup>. The last three formations are all in Western Pennsylvania. Samples for which there were no reliable lithological constraints for the source wells, or for which the analyses indicated significant calcite oversaturation, were not used. The  $p_{\text{CO}_2}$  of each sample was calculated on the basis of its pH and its bicarbonate concentration.

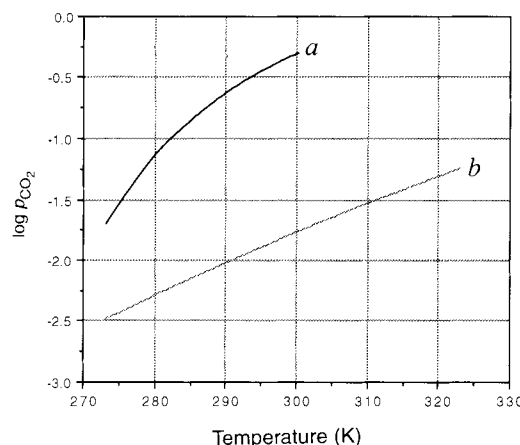


FIG. 3 *a*, Heavy line, top: one-dimensional climate model prediction of  $p_{\text{CO}_2}$  as a function of global mean annual temperature at 2.75 Gyr. *b*, Lower line, estimate of the maximum  $p_{\text{CO}_2}$  in soil waters of siderite-free palaeosols as a function of temperature at  $M_{\text{H}_4\text{SiO}_4} = 10^{-3.4} \text{ M}$ , the highest silica concentration we found in groundwater studies.

enhanced  $\text{CO}_2$  pressure is used to compensate for a solar luminosity that was only  $\sim 80\%$  of its present value<sup>27,28</sup>. Kasting<sup>5,6</sup> has estimated that a  $p_{\text{CO}_2}$  of  $\sim 10^{-0.7}$  atm was needed to maintain today's average surface temperature of  $15^\circ\text{C}$  at 2.75 Gyr. The minimum value of atmospheric  $p_{\text{CO}_2}$  would have been lower if the average surface temperature was lower. However, the scarcity of evidence for late Archaean glacial deposits<sup>3,29</sup> makes the possibility of lower mean annual temperatures unattractive. At average surface temperatures higher than  $15^\circ\text{C}$ , the gap between the atmospheric  $\text{CO}_2$  pressure required to supply sufficient greenhouse forcing and the maximum  $\text{CO}_2$  pressure allowed by the absence of siderite in the lower parts of the palaeosols becomes progressively larger (Fig. 3).

Radiative feedback processes, such as those due to changes in cloud cover and cloud type could have made the Earth's surface temperature much more sensitive to atmospheric  $\text{CO}_2$  content than suggested by cloud-free, one-dimensional climate models. Rossow *et al.*<sup>30</sup> estimated that the inclusion of cloud feedback could increase by as much as  $6^\circ\text{C}$  the temperature predicted by one-dimensional models of an Earth receiving  $80\%$  of today's solar radiation and an atmosphere containing five times the present atmospheric level of  $10^{-3.44}$  atm  $\text{CO}_2$ . Unfortunately, parameterization of cloud feedback is one of the main sources of uncertainty in models of the climate of the near future<sup>31</sup>. Estimates of the effects of cloud feedback on climate at much higher  $\text{CO}_2$  levels and lower solar luminosities should therefore be treated with caution.

If there was insufficient  $\text{CO}_2$  in the atmosphere between 2.75 and 2.2 Gyr to account for the mean surface temperature of the Earth, then other greenhouse gases must also have been present. Sagan and Mullen<sup>32</sup> have suggested that  $\sim 10^{-5}$  atm ammonia could have provided the necessary radiative forcing. However, the sensitivity of  $\text{NH}_3$  to solar ultraviolet radiation makes  $\text{NH}_3$  pressures of  $10^{-5}$  atm difficult to maintain. Also, even at this rather modest  $p_{\text{NH}_3}$ , sea water would have contained so much  $\text{NH}_3$  that the ability of early photosynthesizing organisms to fix  $\text{N}_2$  becomes very difficult to explain (A. H. Knoll, personal communication). Kiehl and Dickinson's suggestion of a methane greenhouse is more appealing<sup>33</sup>. Methane is also ultraviolet labile in the present atmosphere but is much less so in an  $\text{O}_2$ -free atmosphere, especially if the gas is shielded by an organic aerosol smog<sup>34</sup>. Even without such a smog, enough  $\text{CH}_4$  could have been present in an  $\text{O}_2$ -free atmosphere to augment the estimated maximum greenhouse warming due to  $\text{CO}_2$  sufficiently to account for a mean annual temperature

of 15 °C at 2.75-Gyr (J. F. Kasting, personal communication). Buick *et al.*<sup>35</sup> have suggested that an important palaeosol underlies the 3.46-Gyr unconformity below the Warrawoona Group in the Pilbara Craton of Western Australia. This palaeosol could extend the palaeosol-based record for atmos-

pheric  $p_{\text{CO}_2}$  and  $p_{\text{CO}_2}$  by 700 million years. One-dimensional climate models predict that  $p_{\text{CO}_2}$  was substantially higher at 3.46 Gyr than at 2.75 Gyr. The presence or absence of siderite in a pre-Warrawoona palaeosol would serve as an important test of this prediction. □

Received 25 July; accepted 31 October 1995.

- Ohmoto, H. & Felder, R. P. *Nature* **328**, 244–246 (1987).
- Walker, J. G. C. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **82**, 261–289 (1990).
- Towe, K. M. in *Early Life on Earth* (ed. Bengtson, S.) 36–47 (Nobel Symposium 84, Columbia University Press, New York, 1994).
- Hayes, J. M. in *Early Life on Earth* (ed. Bengtson, S.) 220–236 (Nobel Symposium 84, Columbia University Press, New York, 1994).
- Kasting, J. F. *Science* **259**, 920–926 (1993).
- Kasting, J. F. *Precamb. Res.* **34**, 205–229 (1987).
- Button, A. Information Circular 133 (Economic Geology Research Unit, Univ. Witwatersrand, 1979).
- Button, A. & Tyler, N. *Econ. Geol.* **75**, 686–709 (1981).
- Macfarlane, A. W., Danielson, A. & Holland, H. D. *Precamb. Res.* **65**, 297–317 (1994).
- Postma, D. J. *sedim. Petrol.* **47**, 1089–1098 (1977).
- Postma, D. *Chem. Geol.* **31**, 225–244 (1981).
- Sutton, S. J. & Maynard, J. B. *Can. J. Earth Sci.* **30**, 60–76 (1993).
- Klein, C. in *Iron Formation: Facts and Problems* (eds Trendall, A. F. & Morris, R. C.) 417–469 (Developments in Precambrian Geology 6, Elsevier, Amsterdam, 1983).
- Eugster, H. P. & Chou, I.-M. *Econ. Geol.* **68**, 1144–1168 (1973).
- Greenberg, J. & Tomson, M. *Appl. Geochem.* **7**, 185–190 (1992).
- Bruno, J., Wersin, P. & Stumm, W. *Geochim. cosmochim. Acta* **56**, 1149–1155 (1992).
- Cotter, R. D., Young, H. L., Petri, L. R. & Prior, C. H. USGS Water Supply Paper **1759-B-E** (1965).
- Baker, J. A. USGS Water Supply Paper **1257** (1955).
- Wunsch, D. R. *Kentucky Geological Survey Thesis Series* 5, XI (1993).
- Powell, J. D. & Larson, J. D. USGS Water Supply Paper **2274** (1985).

- Poth, C. W. Bulletin W16, 4th series (Pennsylvania Geological Survey, 1963).
- Poth, C. W. Bulletin Water Resources report 36, 4th series (Pennsylvania Geological Survey, 1973).
- Williams, D. R. & McElroy, T. A. USGS Open File Report **90-384** (1991).
- May, F. *Mineralog. Mag.* **58A**, 577–578 (1994).
- Pinto, J. P. & Holland, H. D. *Geol. Soc. Am. spec. Pap.* **216**, 21–34 (1988).
- Arndt, N. T., Nelson, D. R., Compston, W., Trendall, A. F. & Thorne, A. M. *Austr. J. Earth Sci.* **38**, 261–281 (1991).
- Gough, D. O. *Solar Physics* **74**, 21–34 (1981).
- Kasting, J. F. & Grinspoon, D. H. in *The Sun in Time* (eds Sonett, C. P., Giampapa, M. S. & Matthews, M. S.) 447–462 (Univ. Ariz., Tucson, 1991).
- Hoffman, P. F. *Nature* **375**, 537–538 (1995).
- Rossow, W. B., Henderson-Sellers, A. & Weinreich, S. K. *Science* **217**, 1245–1247 (1982).
- Cess, R. D. *et al. J. Geophys. Res.* **95**, 16601–16615 (1990).
- Sagan, C. & Mullen, G. *Science* **177**, 52–56 (1972).
- Kiehl, J. T. & Dickinson, R. E. *J. Geophys. Res.* **92**, 2991–2998 (1987).
- Sagan, C. & Chyba, C. *Abstr. 1st Int. Conf. on Circumstellar Habitable Zones*, 2 (NASA, Ames Research Center, 1994).
- Buick, R. *et al. Nature* **375**, 574–577 (1995).
- Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans* (Princeton Univ. Press, 1984).
- Holland, H. D. & Beukes, N. J. *Am. J. Sci.* **290A**, 1–34 (1990).
- Prasad, N. & Roscoe, S. M. *Geol. Surv. Can. Pap.* **91-1C**, 43–54 (1991).

ACKNOWLEDGEMENTS. We thank J. Kasting for comments and suggestions. This research was supported by NASA.

## An exceptionally widespread ignimbrite with implications for pyroclastic flow emplacement

C. J. N. Wilson\*, B. F. Houghton\*, P. J. J. Kamp† & M. O. McWilliams‡

\* Institute of Geological and Nuclear Sciences, Wairakei Research Centre, Private Bag 2000, Taupo, New Zealand

† Department of Earth Sciences, University of Waikato, Private Bag 3105, Hamilton, New Zealand

‡ Department of Geophysics, Stanford University, Stanford, California 94305, USA

PUMICE-rich pyroclastic flows leave deposits called ignimbrites<sup>1</sup>, which are common in the eruption records of arc and intracontinental volcanoes<sup>2,3</sup>. Pyroclastic flows represent the most dangerous manifestation of volcanism<sup>4</sup> because of their volumes, rapid emplacement at high temperatures, and capacity to extend >100 km from the source. However, maximum distances travelled and extents of areas destroyed by pyroclastic flows remain poorly known. Historic examples have been relatively small, the largest reaching to only ~35 km from the vent<sup>5</sup>, but prehistoric ignimbrites over 150 km across are known<sup>3,6–10</sup>. The largest ignimbrites are invariably partly buried or eroded, and only minimum estimates of their size (and hence of their destructive capacity) can be made. Here we describe the ~1-Myr-old Kidnappers ignimbrite, erupted from the Taupo Volcanic Zone in New Zealand, which we have correlated for ≥385 km and which thus represents the most widespread ignimbrite yet known. We suggest that this wide distribution primarily reflects a coupling of high initial flow velocities with large mass fluxes, and that favourable flow paths and emplacement of flows over a wet substrate may have aided their mobility.

The central Taupo Volcanic Zone (TVZ; Fig. 1) is the most productive Quaternary rhyolitic volcanic system on Earth. At least 34 voluminous ignimbrites (30–1,000 km<sup>3</sup> bulk volume) have been erupted from eight caldera volcanoes<sup>11,12</sup> in central

TVZ since 1.6 Myr. Several of the New Zealand ignimbrites cover areas that imply travel distances of at least 80–100 km (refs 7, 8, 13), distances typical of medium to large ignimbrites elsewhere<sup>3</sup>. However, in all but two examples<sup>7,8</sup>, the maximum distances reached and areas covered by these ignimbrites have not been established with confidence, because the thinner and/or non-welded outer reaches of the deposits have been eroded.

We have studied an ignimbrite (here termed Kidnappers) which, together with its associated earlier fall deposit and its later fluvial and marine reworked products, has previously been described in different parts of the North Island and in deep-sea cores under a variety of names (Table 1; Fig. 1). These deposits, especially the ignimbrite, have not previously been correlated, and so the large eruption magnitude and immensely widespread nature of the ignimbrite have not been recognized before. The Kidnappers ignimbrite is distinguished from all other New Zealand Quaternary ignimbrites by the following features, which allow correlation over large distances:

- (1) completely non-welded texture, although often with pink thermal oxidation colours;
- (2) stratigraphic position above a widespread co-eruptive phreatomagmatic fall deposit;
- (3) common carbonized plant fragments ranging from leaves to 20-cm-diameter logs;
- (4) ferromagnesian mineral assemblage of hypersthene, hornblende and biotite, with the biotite altered to a conspicuous golden hydromica in weathered outcrops;
- (5) <sup>40</sup>Ar–<sup>39</sup>Ar and fission-track age determinations in the range 0.97 to 1.08 Myr (refs 11, 14).
- (6) normal magnetic polarity, interpreted to lie within the (0.99–1.07 Myr; ref. 15) Jaramillo Subchron<sup>14,16</sup>.

This extraordinarily widespread deposit (Fig. 1) is interpreted as an ignimbrite from the presence of lithic- and crystal-enriched segregation bodies, generated by excessive gas fluxes associated with pyrolysis of vegetation. (In contrast, water-reworked mass-flow deposits associated with other ignimbrites known to us lack such features even though texturally they otherwise can be very similar.) The ignimbrite comprises at least two flow units separated, in the northern distal area, by a few centimetres of fall

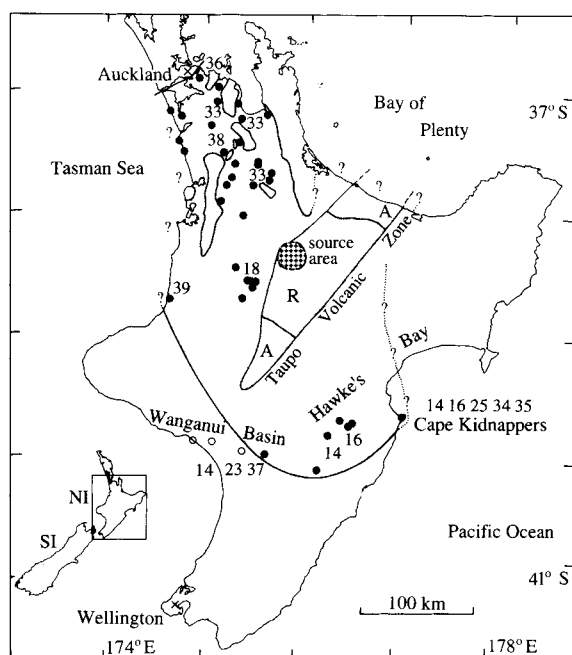


FIG. 1 Map of part of the North Island of New Zealand, showing the minimum inferred original extent of the Kidnappers ignimbrite (shaded area). The east and northeast limit of the Kidnappers ignimbrite is uncertain owing to uplift and erosion, but we infer that the shaded area was covered by the ignimbrite to match the distribution in other sectors. Filled circles represent sites visited by us; open circles are sites where the reworked material described as 'Potaka Tephra' is documented. Numbers refer to references where parts of the Kidnappers ignimbrite and related deposits are previously described from limited geographic areas (Table 1). The outlined area is the Taupo Volcanic Zone, divided into its three segments, dominated by andesitic (A) and rhyolitic (R) volcanism<sup>12</sup> as shown, and the approximate inferred source area for the Kidnappers eruption is marked. Inset (bottom left) shows the map area within New Zealand.

material. The ignimbrite is inferred to have originally covered a minimum area of  $\sim 45,000 \text{ km}^2$ ; observed thicknesses are 1.5–40 m and our best average-thickness estimate is 10 m, yielding a bulk volume of  $450 \text{ km}^3$ . The fall deposit volume is poorly controlled, but comparison with a younger fall deposit that shows comparable thickness variations<sup>17</sup> suggests a bulk volume of several hundred cubic kilometres.

The Kidnappers deposits resulted from a large phreatomagmatic eruption<sup>18</sup>; evidence for this includes the poor sorting, multiple bedding and ash-rich nature of the fall deposit, the presence of accretionary lapilli, and the non-welded nature of the ignimbrite. The eruptive source is not well established, but probably was in the west-central TVZ (Fig. 1) in an area previously identified as a major caldera volcano active in this period<sup>18</sup>. Deep-sea ash cores from the Tasman Sea and southwest Pacific Ocean<sup>19, 21</sup> contain a normally-magnetized major tephra layer (Table 1), indirectly dated from palaeomagnetism to be  $\sim 1.0 \text{ Myr}$  (ref. 15), and previously correlated<sup>22</sup> with one onshore outcrop of the Kidnappers deposits (at Cape Kidnappers; Fig. 1). We propose that the deep-sea tephra is the offshore component of the Kidnappers eruption, possibly augmented by material from eruption of the immediately-succeeding Rocky Hill ignimbrite. The latter<sup>18</sup> shows some geochemical similarities with the Kidnappers ignimbrite (ref. 13 and C.J.N.W., unpublished results), is normally magnetized<sup>14</sup> and has a similar radiometric age<sup>11</sup>. However, the two ignimbrites are separated by a surface representing a period of erosion and weathering<sup>18</sup>, or (in the Wanganui Basin; Fig. 1) by a few metres of pumiceous fluvial sediments. In addition, a volcaniclastic deposit termed the 'Potaka Tephra'<sup>14, 23</sup> has been described from Hawke's Bay and Wanganui Basin and linked<sup>14</sup> with the Rocky Hill ignimbrite. However, our studies show that the 'Potaka Tephra' is a composite deposit that includes both distal Kidnappers and Rocky Hill ignimbrites in the Wanganui Basin, the Kidnappers ignimbrite in Hawke's Bay, and fluvially reworked material from both eruptions in both areas.

An accurate age for the Kidnappers eruption can be obtained by correlating enclosing sediments<sup>16</sup> with O-isotope stages 27–29, which cover the Jaramillo Subchron<sup>15</sup>. This constrains the eruption age to early stage 28; that is, between 1.005 and 1.02 Myr using the astronomical palaeomagnetic chronology<sup>15</sup>. The 'Potaka Tephra' has previously been postulated to correlate with O-isotope stages 27 (ref. 14) or 28 (ref. 23), but note that,

depending on whether the material came from the later Rocky Hill or earlier Kidnappers eruptions, respectively, both interpretations may be correct but at different sites.

Preserved outcrops of the Kidnappers ignimbrite span a bilaterally symmetric area 385 km across centred on the inferred source area in the TVZ (Fig. 1). Why did the Kidnappers ignimbrite reach such large distances, and are the factors involved general enough to be useful in understanding all such far-travelled pyroclastic flows? We identify three key factors here: the first is specific, the second and third are more general.

First, the Kidnappers flow paths were relatively favourable, certainly to the north, and probably also to the south and west. Once the flows had left the immediate confines of the source volcano, there were few large ( $>300 \text{ m}$ ) topographic barriers to the north and most of the ignimbrite was deposited on a continuing gentle ( $<1^\circ$ ) downhill gradient from the source. Modern topographic barriers just south of Auckland City represent terrain that has mostly been elevated by block faulting and mafic volcanism since 1.0 Myr (ref. 24). To the south and west, present-day mountains present a barrier to all but the most violently emplaced pyroclastic flows, but field evidence implies that major uplift and dissection of these mountains began after the Kidnappers eruption (and continue to the present day<sup>14, 18, 25</sup>).

Second, in the Kidnappers eruption the underlying wet phreatomagmatic fall deposits could have acted as sources of gas, maintaining fluidity and mobility of the pyroclastic flows. At one section in the Kidnappers ignimbrite where the key evidence has been preserved, elongate masses of fall material occur streaming upwards and disaggregating into the ignimbrite, accompanied by abundant segregation pipes, implying that heat from the flow had boiled interstitial water in the fall deposits. The overall effect on the fluidity and mobility of the flow can therefore be likened to pouring water into a hot skillet. Thus the average depositional slope on the Kidnappers ignimbrite in northern areas that have undergone minimal subsequent tilting is  $\leq 0.02^\circ$ , compared with typical aggradational slopes on New Zealand Quaternary welded ignimbrites of  $0.5\text{--}2^\circ$ .

Third, the distances reached by pyroclastic flows in general are dependent on both their initial velocities and masses. One extreme of the resulting deposits are high-aspect-ratio ignimbrites<sup>1</sup> (aspect ratio (a.r.) is the ratio of average thickness of a deposit to the diameter of a circle that covers the same area), with a.r. values of 1: ( $<1,000$ ), whose flow termini rep-



TABLE 1 Compendium of previously published occurrences and names applied to Kidnappers eruption deposits

| Previously applied name | Nature of deposits recognized      | Area relative to source | Reference* |
|-------------------------|------------------------------------|-------------------------|------------|
| Potaka Tephra           | Volcaniclastic sediments           | S, Se                   | 14, 23, 37 |
| Kidnappers Ignimbrite   | Ignimbrite over fall deposit       | SE                      | 16, 25, 34 |
| Unit E                  | Ignimbrite over fall deposit       | W                       | 18         |
| Kidnappers Tuff         | Volcaniclastic sediments           | SE                      | 22, 35     |
| Morrinsville ignimbrite | Ignimbrite over fall deposit       | N                       | 33         |
| (No name)               | Ignimbrite                         | N, W                    | 36, 38, 39 |
| Ash A                   | Primary tephra fall on ocean floor | SW Pacific              | 19         |
| Ash M5                  | Primary tephra fall on ocean floor | SW Pacific              | 20         |
| Ash T8 (Hole 592)       | Primary tephra fall on ocean floor | Tasman Sea              | 21         |
| Ash T9 (Hole 593)       | Primary tephra fall on ocean floor |                         |            |

\* Numbers also refer to areas marked on Figure 1.

resent points where the flows simply slowed to a halt. In contrast, in the best documented low-aspect-ratio ignimbrite<sup>7,26</sup>, with an a.r. of 1:105,000, the ignimbrite termini represent points where, although the flow was still moving, it had run out of material to deposit. In the latter case, inferred eruption rates and initial velocities were extremely high<sup>26</sup>, but the volumes of material erupted were inadequate to generate far-travelled flows. We suggest that the Kidnappers ignimbrite (a.r. 1:24,000) represents an efficient compromise whereby large volumes were erupted at sustained high rates, providing both a high initial velocity and the volume to keep travelling and depositing for >190 km and still not exhaust the supply of material.

Two other low-aspect-ratio ignimbrites that approach the extremely wide dispersal of the Kidnappers ignimbrite are the 640-km<sup>3</sup> Miocene Peach Springs ignimbrite (USA)<sup>9</sup> and the 330-km<sup>3</sup> Miocene Rattlesnake ignimbrite (USA)<sup>10</sup>, which have a.r. values of 1:10,000, and between 1:4,000 and 1:20,000, respectively. Of the factors outlined above in assessing the mobility of the Kidnappers ignimbrite, only the idea that there is an 'optimum' combination of mass and velocity for maximum travel distance will hold for these other examples, neither of which was phreatomagmatic or had a notably favourable emplacement path.

With moderate to large volume ignimbrites there is a general contrast between those that are of intermediate to low aspect-ratio (a.r. 1:(>1,000)), and show a radial or bilateral symmetri-

cal distribution about the source area (for example, Kidnappers, Oruanui<sup>8</sup>, Peach Springs<sup>9</sup> and Rattlesnake<sup>10</sup>), and those that are less far-travelled, have a high aspect ratio, and are noticeably asymmetrically distributed (for example, Matahina<sup>27</sup> (a.r. 1:850), Kaingaroa<sup>28</sup> (a.r. 1:650) and Bishop<sup>29</sup> (a.r. 1:300)). In the Bishop, the ignimbrite asymmetry is attributed to its eruption from widely spaced multiple vents along a caldera ring fracture<sup>30</sup>. These multiple vents would also have acted to reduce the exit velocity of the outflowing ejecta for a given eruption rate, and hence reduced the initial impetus given to the flows. In contrast, in the Oruanui eruption the ignimbrite is interpreted to have come from a single source now within the caldera<sup>8</sup> (relevant information is lacking for the Peach Springs and Rattlesnake ignimbrites).

We thus speculate that, for a given eruption rate, the style of vent development during eruption (single vent versus multiple, ring-fracture-controlled vents) will influence how rapid an initial impetus is given to the flows, and hence how far a flow of a given mass or volume will travel. Furthermore we suggest that the most widespread ignimbrites are those that balance high mobility with substantial volume, with an optimum bulk volume of a few hundred cubic kilometres. Ignimbrites of larger volume (for example, Youngest Toba Tuff<sup>6</sup>; area 20,000 km<sup>2</sup>, volume 2,000 km<sup>3</sup>, a.r. 1:1,600) tend to be no more widespread, possibly because the processes of syn-eruptive caldera collapse and concomitant opening of multiple vents<sup>31</sup> lead to more sluggish discharges (even though the overall eruption rate may have greatly increased) and the build-up of thicker deposits. Ignimbrites of smaller volume but emplaced at high velocities like Taupo<sup>7</sup> (area 20,000 km<sup>2</sup>, volume 30 km<sup>3</sup>, a.r. 1:105,000) in turn lack the material to take full advantage of the high velocities imparted to the flow(s) at source.

Thus in modelling pyroclastic flow eruptions and predicting the distance to which the largest pyroclastic flows will travel, we infer that account has to be taken not only of the eruption volume but also the style of caldera collapse. In this regard, the overall eruption rate seems to be less important than the ratio of vent area to eruption rate, with its link to ejecta velocities and hence flow velocities<sup>32</sup>. We speculate that, for a given eruption rate, a single vent will promote higher ejecta velocities and symmetrically distributed ignimbrites, whereas if multiple vents develop, the ejecta and hence flow velocities will be reduced and the flows directed asymmetrically. Other factors such as a favourable pathway and the phreatomagmatic nature of an eruption, which were important in enabling the Kidnappers ignimbrite to cover such a large area, seem to be absent or to play secondary roles in pyroclastic flows in general, as such factors were absent from the Peach Springs and Rattlesnake ignimbrites. □

Received 19 July; accepted 23 October 1995.

- Walker, G. P. L. *J. Volcan. geotherm. Res.* **17**, 65–88 (1983).
- Ross, C. S. & Smith, R. L. *Prof. Pap. U.S. geol. Surv.* **366**, 1–81 (1960).
- Cas, R. A. F. & Wright, J. V. *Volcanic Successions Modern and Ancient* (Allen & Unwin, London, 1987).
- Walker, G. P. L. *Interdisc. Sci. Rev.* **7**, 148–157 (1982).
- Self, S., Rampino, M. R., Newton, M. & Wolff, J. A. *Geology* **12**, 659–663 (1984).
- Aldiss, D. T. & Ghazali, S. A. *J. geol. Soc. Lond.* **141**, 487–500 (1984).
- Wilson, C. J. N. *Phil. Trans. R. Soc. Lond. A* **314**, 229–310 (1985).
- Wilson, C. J. N. *Bull. volcan.* **53**, 635–644 (1991).
- Buesch, D. C. *Bull. U.S. geol. Surv.* **2053**, 55–85 (1993).
- Streck, M. J. & Grunder, A. L. *Bull. volcan.* **57**, 151–169 (1995).
- Houghton, B. F. *et al. Geology* **23**, 13–16 (1995).
- Wilson, C. J. N. *et al. J. Volcan. geotherm. Res.* **68**, 1–28 (1995).
- Briggs, R. M. *et al. J. Volcan. geotherm. Res.* **56**, 175–203 (1993).
- Shane, P. A. N. *J. Geol. Geophys. Res.* **37**, 25–36 (1994).
- Shackleton, N. J., Berger, A. & Peltier, W. R. *Trans. R. Soc. Edinb.* **81**, 251–261 (1990).
- Kamp, P. J. J. *Misc. Pub. geol. Soc. N.Z.* **50B**, 105–118 (1990).
- Wilson, C. J. N. *Bull. U.S. geol. Surv.* **2047**, 93–99 (1994).
- Wilson, C. J. N. *Bull. R. Soc. N.Z.* **23**, 179–193 (1986).
- Ninkovich, D. *Earth planet. Sci. Lett.* **4**, 89–102 (1968).
- Watkins, N. D. & Huang, T. C. *N.Z. J. Geol. Geophys.* **20**, 179–198 (1977).
- Nelson, C. S., Froggatt, P. C. & Gosson, G. J. *Init. Rep. DSDP 1161–1172* (1988).
- Seward, D. N. *J. Geol. Geophys. Res.* **18**, 507–510 (1975).
- Pillans, B. J., Roberts, A. P., Wilson, G. S., Abbott, S. T. & Alloway, B. V. *Earth planet. Sci. Lett.* **121**, 81–98 (1994).

- Briggs, R. M., Okada, T., Itaya, T., Shibuya, H. & Smith, I. E. M. *N.Z. J. Geol. Geophys.* **37**, 143–153 (1994).
- Kamp, P. J. J. in *Landforms of New Zealand* (eds Soons, J. M. & Selby, M. J.) 233–254 (Longman Paul, Auckland, 1982).
- Wilson, C. J. N. & Walker, G. P. L. in *Tephra Studies* (eds Self, S. & Sparks, R. S. J.) 441–448 (Reidel, Dordrecht, 1981).
- Bailey, R. A. & Carr, R. G. *N.Z. J. Geol. Geophys.* **37**, 319–344 (1994).
- Nairn, I. A., Wood, C. P. & Bailey, R. A. *Bull. volcan.* **56**, 529–537 (1994).
- Hildreth, W. *Spec. Pap. geol. Soc. Am.* **180**, 43–75 (1979).
- Hildreth, W. & Mahood, G. A. *Bull. geol. Soc. Am.* **97**, 396–403 (1986).
- Druitt, T. H. & Sparks, R. S. J. *Nature* **310**, 679–681 (1984).
- Sparks, R. S. J., Wilson, L. & Hulme, G. J. *geophys. Res.* **83**, 1727–1739 (1978).
- Walker, G. P. L. & Wilson, C. J. N. *J. Volcan. geotherm. Res.* **18**, 117–133 (1983).
- Kamp, P. J. J. in *Proceedings of Tephra Workshop, June 1980* (eds Howarth, R., Froggatt, P. C., Vucetich, C. G. & Collen, J. D.) 52–54 (Publ. 20, Dept. Geol., Victoria Univ. Wellington, 1981).
- Kingma, J. T. *Bull. N.Z. geol. Surv.* no. 70 (Govt Printer, Wellington, 1971).
- Kermode, L. O. *Inst. Geol. Nucl. Sci., Geol. Map* no. 2 (1992).
- Alloway, B. V., Pillans, B. J., Sandhu, A. S. & Westgate, J. A. *Sedim. Geol.* **82**, 299–310 (1993).
- Nelson, C. S., Mildenhall, D. C., Todd, A. J. & Pocknall, D. T. *N.Z. J. Geol. Geophys.* **31**, 21–40 (1988).
- Duff, R. O. *Misc. Pub. geol. Soc. N.Z.* **80A**, 58 (1994).

ACKNOWLEDGEMENTS. We thank J. D. L. White and S. Self for comments. This work was supported by the N.Z. Foundation for Research Science & Technology (C.J.N.W., B.F.H.), the University of Waikato Research Committee (P.J.J.K.) and the Royal Society (C.J.N.W.).

# Possible flush instability in mantle convection at the Archaean-Proterozoic transition

D. Breuer & T. Spohn

Institut für Planetologie, Wilhelm-Klemm-Strasse 10,  
48149 Münster, Germany

THE late Archaean eon (3 to 2.5 billion years ago) seems to have been a time of profound geological changes<sup>1-5</sup>: rapid growth of the continents, of the volume of cratonic sediments and of the area of the continental shelves was accompanied by widespread emplacement of potassium-rich granitic rocks, leading to the 'cratonization' of the continental crust<sup>3,6</sup>. There is also evidence for climate change<sup>7,8</sup> at this time, possibly reflecting changes in the composition of the atmosphere and the volume of the oceans. The period ends with the Huronian glaciation, the first well documented glacial episode<sup>9</sup>, and this transition appears to be marked by a strong increase in the intensity of the geomagnetic field<sup>10</sup>. Here we present simulations of mantle convection at the end of the late Archaean, and show that a breakdown of two-layer convection (a 'flush' instability<sup>11</sup>) occurring at this time could have led to a period of volcanic and tectonic activity that might account for the reported geological and palaeoenvironmental changes.

Recent calculations suggest that the spinel-perovskite phase transition at 670 km depth, the boundary between the upper and lower mantles, may play a crucial role in the dynamics of the mantle<sup>11,12</sup>. The phase boundary appears to stabilize two-layer convection at high Rayleigh numbers, whereas at low Rayleigh numbers ( $\leq 10^7$ ; refs 13, 14) convection currents easily penetrate the boundary and the flow comprises the whole mantle. When

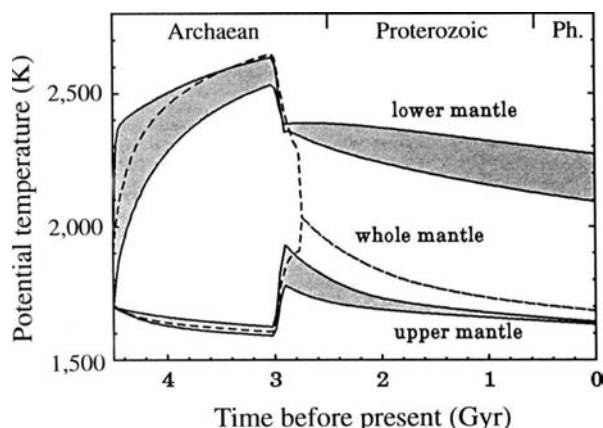


FIG. 1 The potential temperatures of the upper and lower mantles and the whole mantle as functions of time before present. The potential temperature is the average temperature extrapolated adiabatically to surface pressure. The dashed line refers to the model with whole-mantle convection following the flush instability. The shaded area brackets the results obtained for the model with intermittent mass transfer between the layers following the instability and with the minimum value of the mass flow rate between the mantle layers considered. For these models the following parameters have been varied: the thermal conductivity of the lower mantle,  $4\text{--}16\text{ W m}^{-1}\text{ K}^{-1}$ ; the activation temperature for creep of the lower mantle,  $(5\text{--}7) \times 10^4\text{ K}$ ; the initial potential lower mantle temperature,  $1,700\text{--}1,900\text{ K}$ ; and the factor of increase of the heat transfer rate from the lower mantle during the instability,  $15\text{--}30$ . Other important model parameter values are  $4\text{ W m}^{-1}\text{ K}^{-1}$  for the upper-mantle thermal conductivity,  $5 \times 10^4\text{ K}$  for the upper-mantle activation temperature for creep,  $8 \times 10^6\text{ J m}^{-3}\text{ K}^{-1}$  for the mantle heat capacity, and  $0.3$  for the exponent in the Nusselt-Rayleigh number relation. Ph, Phanerozoic.

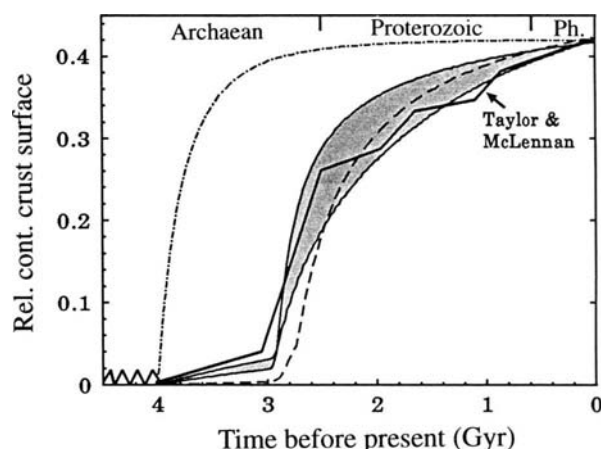


FIG. 2 Area of the continental crust as derived from geochemical data and a constant freeboard in the post-Archaean by Taylor and McLennan<sup>2,3</sup> (thick line) compared with the results of evolution calculations. The solid lines and shaded area refer to the model with intermittent mass transfer between the mantle layers following the flush instability and with the minimum value of the mass flow rate between the mantle layers considered (compare Fig. 1); the dashed line refers to the model with whole mantle convection after the instability. For further comparison, the crust growth curve calculated for a model with whole-mantle convection throughout the entire evolution is shown (dash-dotted line)<sup>19</sup>. Formation of the continental crust is assumed to start 4 Gyr ago, after the end of the post-accretional bombardment. The thickness of the continental crust is 37 km and the present heat flow into the base of the crust is  $35\text{ mW m}^{-2}$ . For further explanations, see Fig. 1 legend.

the transition occurs in a cooling mantle whose Rayleigh number decreases with time, cold upper-mantle material breaks through the boundary and flushes into the lower mantle while hot lower-mantle material shoots up into the upper mantle. Thereafter, and before whole-mantle convection is established, layered convection with intermittent mass transfer between the layers triggered by cold avalanches<sup>15</sup> that penetrate the boundary is possible. These instabilities are much smaller in amplitudes of mass and heat flux than the first flush event<sup>13</sup>. Seismic<sup>12</sup> and geochemical<sup>16</sup> evidence suggests that the Earth at present is in the intermittent regime.

Our proposal to associate the Archaean to Proterozoic transition with a major flush event is motivated by the results of a parametrized thermo-chemical evolution model that incorporates these changes in convection layering. The model, which assumes plate tectonics since the early Archaean<sup>17</sup>, starts with stable two-layer mantle convection established soon after the accretion of the planet. The layers are thermally coupled through a thermal boundary layer, but there is no mass exchange between the layers at that time. We parametrize two-layer mantle convection<sup>18</sup>, and include both a simple model of mantle differentiation by the production of the continental crust<sup>19</sup> and a simple core dynamo model<sup>20</sup>.

The rate of growth of continental crust is calculated from the difference between the volumetric crust production and recycling rates, with the following assumptions<sup>19</sup>. The production rate is proportional to the temperature-dependent convective mass flow rate and the concentration of crustal component in the upper mantle. The recycling rate is proportional to the area of the continental crust and the plate speed. The thickness of the continental crust is constant in time. The heat flow into the bottom of the crust is proportional to the crust production rate and is typically a few tens of per cent of the oceanic heat flow. The model fractionates mantle heat sources into the crust at a rate that is proportional to the net crust growth rate. The continental freeboard, the mean elevation of the continents above sea level, was calculated from the thermo-chemical evolution model

through its dependence on the surface of the continents, mantle heat flow, ocean volume, and the thickness of the oceanic crust<sup>21,22</sup>. The ocean volume changes with mantle degassing and regassing. The net degassing rate was taken to be proportional to the net crust growth rate because both depend on mantle convection vigour in a similar way<sup>23</sup>.

Crucial parameters are the activation temperature for creep in the lower mantle and the lower-mantle thermal conductivity. These values were varied over a wide range. The constants of proportionality for the crust production and recycling rates were adjusted to fit the present volume of the crust and its mean age<sup>19</sup>. The former constant should actually vary with temperature and degree of partial melting in the mantle. For simplicity, this has been neglected which results in a small underestimate of crust growth during the instability. The constant of proportionality for net mantle degassing was adjusted to fit the present volume of the oceans. Other parameter values were chosen as in previous work<sup>18,20</sup>.

After 1.5 Gyr of evolution, approximately the beginning of rapid crust growth<sup>2</sup>, the flush instability was assumed to occur. The Rayleigh number at that time is within the range of uncertainty of the critical Rayleigh number for the onset of the instability<sup>14</sup>. It was simulated by allowing a sufficiently rapid mass and heat exchange between the mantle layers for a period of 100 Myr (ref. 13) such that 50% of the volume of continental crust is produced by the end of the Archaean. The heat transfer rate from the lower mantle was increased by a factor varied between 15 and 30 which is reasonably of the order of the lower-mantle Nusselt number. For the evolution after the instability,

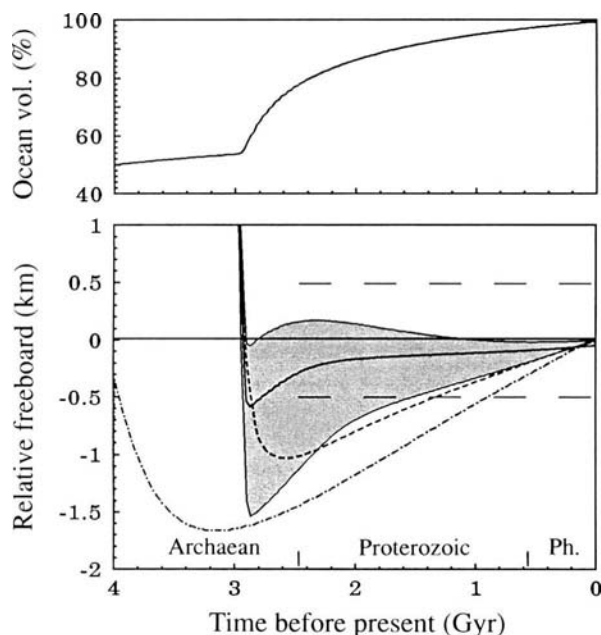


FIG. 3 Continental freeboard relative to the present value (bottom panel) and ocean water volume (top panel) in per cent as functions of time before present. The dashed horizontal lines bracket the variation of the freeboard ( $\pm 500$  m) in the post-Archaeon as suggested by observation<sup>22</sup>. The solid lines in both panels refer to our preferred model with intermittent mass transfer between the mantle layers following the flush instability. Parameter values for this model are:  $4 \text{ W m}^{-1} \text{ K}^{-1}$  for the lower-mantle thermal conductivity,  $5 \times 10^4$  for the activation temperature for lower-mantle creep,  $1,900 \text{ K}$  for the initial lower-mantle temperature, and 15 for the factor of increase of the heat transfer rate from the lower mantle during the instability. The shaded area indicates the results of parameter variations (compare Fig. 1). The dashed line refers to the model with whole-mantle convection after the instability. For further comparison we show the freeboard calculated for a model with whole-mantle convection throughout the entire evolution<sup>19</sup> (dash-dotted line). For further explanations see Fig. 1.

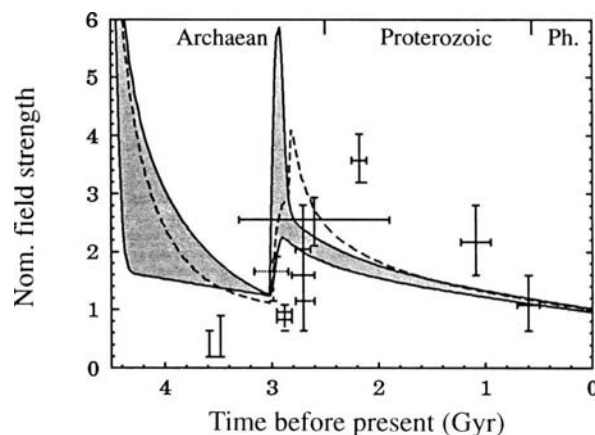


FIG. 4 Nominal magnetic field strength scaled to the present-day value, calculated from a power balance of the core according to the method outlined in ref. 20 and compared with data (points with error bars) in ref. 10. For further explanations see Fig. 1.

we considered a wide range of mass transfer rates between the mantle layers, from  $\sim 10\%$  of the average lower-mantle convection mass flow rate to whole-mantle convection. The heat transfer rate from the lower mantle in the minimum model is larger by a factor of two than the conductive heat flow through the boundary layer.

Results in terms of potential mantle temperature, continental crust area and freeboard, ocean volume and magnetic field strength are shown in Figs 1–4. Before the flush instability, the lower mantle heats up in two-layer convection (Fig. 1); it cools by  $200\text{--}600 \text{ K}$  through the instability while the upper mantle heats up by  $200\text{--}400 \text{ K}$ . The major pulse of crust growth (Fig. 2) at the transition is caused by the increase in upper-mantle convection vigour and temperature and by the lower-mantle reservoir becoming available for crust growth. A model based entirely on whole-mantle convection<sup>19</sup> is found to be inconsistent with the empirical growth curve although it is consistent with an alternative growth model<sup>1,24</sup>. The continental freeboard (Fig. 3) decreases significantly during the instability and remains within the error bounds of the geological data in the post-Archaeon. The model suggests that about half of the ocean volume degassed during and after the instability. The evolution of the magnetic field (Fig. 4) shows a significant increase in the dipole moment. For simplicity, we have assumed a thermal dynamo, and have neglected a possible chemically driven dynamo which is more efficient at converting potential energy into magnetic field energy than the thermal dynamo is at converting thermal energy. We note, however, that our main point is valid in both cases: the rapid cooling of the lower mantle through the instability will enhance core cooling and thereby the vigour of convection in the core. This should have major effects on the dynamo.

The growth curve of the continental crust, the freeboard, and the magnetic field strength are all remarkably consistent with the geological data. A systematic search of parameter space has shown that these results are insensitive both to initial conditions and to variations of parameter values over a wide range. The increase in upper-mantle potential temperature by  $200\text{--}400 \text{ K}$  and in average heat flow by a factor of three provides an explanation for the proposed<sup>3,6</sup> cratonization of continental crust by intracrustal differentiation through partial melting of the lower crust. The breakdown of two-layer mantle convection is likely to have been accompanied by a major change in the horizontal dimensions of surface plates from several hundred kilometres, the depth scale of the upper mantle, to a few thousand kilometres. It has been suggested that the Archaean crust was characterized by a large number of microplates and that island-arc

magmas originated in relatively warm basaltic crust at lower pressures than at present<sup>3</sup>. Simple scaling shows that, at the same potential mantle temperature, the subducting oceanic crust would have been a few hundred degrees warmer if the Archaean plate scale was of the order of the upper-mantle thickness. A low potential temperature of the early to mid-Archaean upper mantle is consistent with the results of thermobarometry<sup>25</sup>. MgO-rich Archaean komatiites<sup>3,26</sup> may have originated in the relatively hot boundary layer between the mantle layers. This boundary layer is significantly cooler or even non-existent after

the transition, thereby explaining the absence of these komatiites in the post-Archaean. The increase in the volume of the sediments and the decrease in atmospheric temperature at the transition may be related to the growth of the continents: an increase in continental surface area increases the weathering rate and the rate of removal of the greenhouse-gas CO<sub>2</sub> from the atmosphere<sup>27</sup>. The formation of the shelves may indicate a dramatic drop in continental freeboard which may, at least partly, be attributed to a substantial increase in ocean water mass through increased degassing of the mantle. □

Received 23 February; accepted 7 November 1995.

1. Armstrong, R. L. *Austr. J. Earth Sci.* **38**, 613–630 (1991).
2. Taylor, S. R. & McLennan, S. M. *The Continental Crust: its Composition and Evolution* (Blackwell Scientific, Boston, 1985).
3. Taylor, S. R. & McLennan, S. M. *Rev. Geophys.* **33**, 241–265 (1995).
4. Kröner, A. *Precambrian Plate Tectonics* (Elsevier, Amsterdam, 1981).
5. Condie, K. C. *Archaean Crustal Evolution* (Elsevier, Amsterdam, 1984).
6. Percival, J. A. in *Archaean Crustal Evolution* (ed. Condie, K. C.) 357–410 (Elsevier, Amsterdam, 1994).
7. Kasting, J. F. *Science* **259**, 920–926 (1993).
8. Des Marais, D. J. in *Archaean Crustal Evolution* (ed. Condie, K. C.) 505–523 (Elsevier, Amsterdam, 1994).
9. Crowley, T. J. *Rev. Geophys. Space Phys.* **21**, 828–877 (1983).
10. Hale, C. J. *Nature* **329**, 233–237 (1986).
11. Christensen, U. & Yuen, D. A. *J. geophys. Res.* **90**, 10291–10300 (1985).
12. Tackley, P. J. *Rev. Geophys.* **33** Suppl., 275–282 (1995).
13. Steinbach, V., Yuen, D. A. & Zhao, W. *Geophys. Res. Lett.* **20**, 1119–1122 (1993).
14. Yuen, D. A. et al. *Phys. Earth planet Inter.* **86**, 185–203 (1994).

15. Solheim, L. P. & Peltier, W. R. *J. geophys. Res.* **99**, 6997–7018 (1994).
16. Stein, M. & Hofman, A. W. *Nature* **372**, 63–68 (1994).
17. Calvert, A. J., Sawyer, E. W., Davis, W. J. & Ludden, J. N. *Nature* **375**, 670–674 (1995).
18. Spohn, T. & Schubert, G. *J. geophys. Res.* **87**, 4682–4696 (1982).
19. Spohn, T. & Breuer, D. in *Evolution of the Earth and Planets* (eds Takahashi, E., Jeanloz, R. & Rubie, D. C.) 55–71 (Geophys. Monogr. Ser. 74, Am. Geophys. Union, Washington DC, 1993).
20. Stevenson, D. J., Spohn, T. & Schubert, G. *Icarus* **54**, 466–489 (1983).
21. Schubert, G. & Reymier, A. P. S. *Nature* **316**, 336–339 (1985).
22. Galer, S. J. G. *Earth planet. Sci. Lett.* **105**, 214–228 (1993).
23. McGovern, P. & Schubert, G. *Earth planet. Sci. Lett.* **96**, 27–37 (1989).
24. Armstrong, R. L. *Phil. Trans. R. Soc. Lond. A* **301**, 433–471 (1981).
25. Boyd, F. P. & Gurney, J. J. *Science* **232**, 427–477 (1986).
26. Arndt, N. T. in *Archaean Crustal Evolution* (ed. Condie, K. C.) 11–44 (Elsevier, Amsterdam, 1994).
27. Tajika, E. & Matsui, T. in *Origin of the Earth* (eds Newsom, H. E. & Jones, J. H.) 347–370 (Oxford Univ. Press, 1990).

ACKNOWLEDGEMENTS. This work was supported by the Deutsche Forschungsgemeinschaft.

## Unexpected dominance of high frequencies in chaotic nonlinear population models

Joel E. Cohen

Rockefeller University, 1230 York Avenue, New York, New York 10021-6399, USA; and NERC Centre for Population Biology, Imperial College, Silwood Park, Ascot SL5 7PY, UK

BECAUSE water has a higher heat capacity than air, large bodies of water fluctuate in temperature more slowly than does the atmosphere<sup>1</sup>. Marine temperature time series are 'redder' than atmospheric temperature time series by analogy to light: in red light, low-frequency variability has greater amplitude than high-frequency variability, whereas in white light all frequencies have the same amplitude<sup>2</sup>. Differences in the relative importance of high- and low-frequency variability in different habitats affect the population dynamics of individual species and the structure of ecological communities<sup>3–9</sup>. Population dynamics of individual species are thought to be dominated by low-frequency fluctuations, that is, to display reddened fluctuations<sup>10</sup>. Here I report, however, that in eight nonlinear, iterative, deterministic, autonomous, discrete-time population models, some of which have been used to model real biological populations, the power spectral densities of chaotic trajectories are neither white nor reddened but are notably blue, with increasing power at higher frequencies.

The simplest discrete-time population model capable of representing the influence of environmental variability assumes that  $P_{t+1} = r_t P_t$ ,  $P_0 > 0$  (ref. 11). Here  $P_t$  denotes the population size (assumed to vary continuously) at time  $t$ , and  $r_t$  is a positive-valued random variable that describes the effect of environmental variation on survival and reproduction from time  $t$  to time  $t+1$ . In this idealized model, the population transforms a white-noise environmental input to a random-walk population trajectory, which has a reddened spectrum. To see this, suppose that  $r_t$  is identically and independently distributed for all  $t$ ;  $r_t$  is thus a discrete-time white noise. Its spectrum is flat, as is shown approximately in the numerical simulations in Fig. 1a. The loga-

rithm of population size  $\log(P_t) = \log(P_0) + \sum_{i=0}^{t-1} \log(r_i)$  is a random walk or a discrete-time brownian motion (sometimes called brown noise), which is known to be highly reddened. The spectrum of a discrete-time random walk is illustrated in the numerical simulations in Fig. 1b. It would hardly be surprising if, in more realistic models or in reality, white or reddened environmental variability resulted in still redder population fluctuations.

I examined the spectra of chaotic trajectories in eight deterministic, autonomous, nonlinear population models (Table 1). In these models,  $P_t$  uniquely determines  $P_{t-1}$ , hence population fluctuations are generated by internal dynamics rather than by environmental variability. The first six models are well known and have been widely used for non-human biological populations with discrete generations<sup>12,13</sup>. The last two models<sup>14,15</sup>, developed to describe human populations, have not previously been shown to behave chaotically for some parameter values. It appears not to have been noticed previously that chaotic behaviour in all eight of these models has a blue spectrum for many parameter values.

For each model, I generated 100 sample paths of 1,024 time steps, starting from an initial population size  $P_0$  which was randomly and uniformly distributed between 0 and 1. (The Austin-Brewer model<sup>14</sup> was exceptional in that  $P_0$  was uniformly distributed between 0 and 20.) I then truncated  $P_0$  through  $P_{512}$  to eliminate any transient effects of initial conditions and checked that no population sizes were zero or negative. The remaining sequence of 512 population sizes was then tested against a sufficient condition for the existence of chaos<sup>16</sup>. For the parameter values shown in Table 1, all simulations satisfied the sufficient condition and were therefore shown to be chaotic. Next I calculated the fast Fourier transform (FFT) of the 512 population sizes, using the FFT function provided by Matlab version 3.5g. Because of the mirror symmetry in the spectrum with respect to its middle, I retained only the first 256 elements returned by the FFT function, dropped the first of these because it was artefactually distorted by the finite length of the time series, and multiplied each of the remaining 255 complex numbers by its complex conjugate to obtain the squared amplitude or power of the spectrum at the corresponding frequency. At each of the 255 frequencies, I computed the minimum, the maxi-



FIG. 1 Power spectral densities (spectra) of discrete-time random sequences: a, white noise  $X_t = \varepsilon_t$ , and b, red noise (or random walk or brown noise)  $X_{t+1} = X_t + \varepsilon_t$ . In both cases,  $\varepsilon_t$  is a sequence of independent and identically distributed normal random variates with mean 0 and variance 1. Sequences of length 512 were generated starting from  $X_0 = 0$  and standardized to have overall mean 0 and variance 1. The fast Fourier transform and spectrum were then computed as described in the text. The continuous lines represent the average spectrum from 100 simulations; the rectangles represent the maximum spectrum; and the triangles represent the minimum spectrum. The horizontal coordinate  $x$  is frequency; the corresponding period is  $1/x$ . For example, the abscissa 0.4 corresponds to 0.4 cycles per time unit or a period of  $1/0.4 = 2.5$  time units.

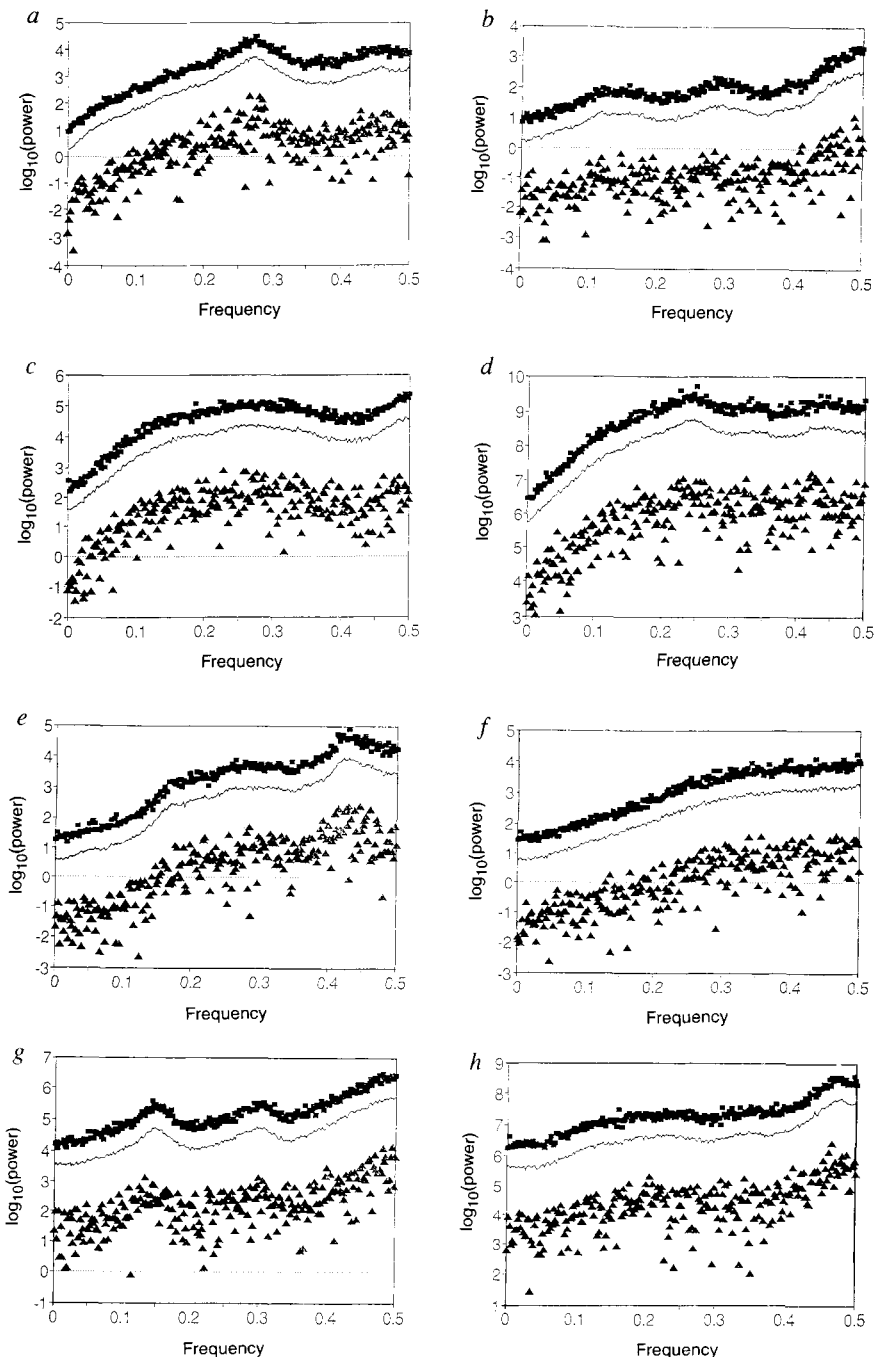
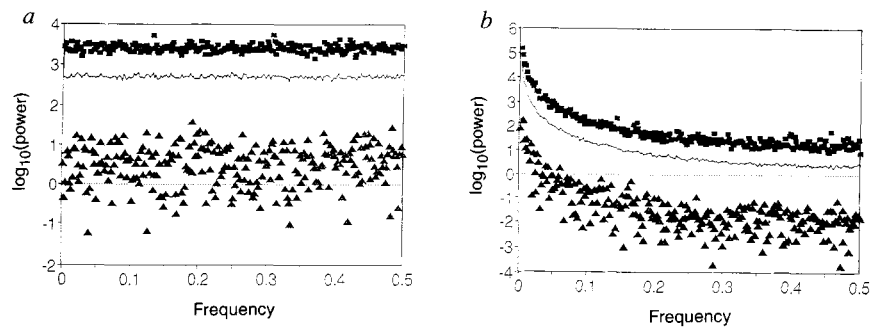


FIG. 2 Power spectral densities of eight discrete-time nonlinear population models: a, Moran-Ricker<sup>12,25,26</sup>; b, Verhulst<sup>12</sup>; c, Pennycuick<sup>12,27</sup>; d, Hassell<sup>12,28</sup>; e, Maynard Smith<sup>12,29</sup>; f, Varley<sup>12,30</sup>; g, Austin-Brewer<sup>14</sup>; and h, Malthus-Condorcet-Mill<sup>15</sup>. Table 1 gives the equations and parameter values. Symbols and units are as in Fig. 1. By contrast with the white or reddened spectra commonly observed in environmental and ecological time series, these spectra are blue, with greater power at higher frequencies.

TABLE 1 Eight nonlinear population models, with parameter values that give chaotic behaviour

| Model                                                                  | Parameter values*             | Model name                           |
|------------------------------------------------------------------------|-------------------------------|--------------------------------------|
| $P_{t+1} = P_t \exp[r(1 - P_t)]$                                       | $r = 3.7$                     | Moran-Ricker <sup>12,25,26</sup>     |
| $P_{t+1} = P_t[1 + r(1 - P_t)]$                                        | $r = 2.7$                     | Verhulst <sup>1,2</sup>              |
| $P_{t+1} = rP_t / \{1 + \exp[-a(1 - P_t/b)]\}$                         | $r = 50, a = 0.1, b = 0.1$    | Pennycuik <sup>12,27</sup>           |
| $P_{t+1} = rP_t / (1 + aP_t)^b$                                        | $r = 55, a = 0.0001, b = 100$ | Hassell <sup>12,28</sup>             |
| $P_{t+1} = rP_t / [1 + (aP_t)^b]$                                      | $r = 5, a = 0.5, b = 4$       | Maynard Smith <sup>12,29</sup>       |
| $P_{t+1} = rP_t$ if $P_t \leq C$ , $P_{t+1} = rP_t^{1-b}$ if $P_t > C$ | $r = 4, b = 3, C = 1$         | Varley <sup>12,30</sup>              |
| $P_{t+1} = P_t \{1 + r[1 - \exp(-sP_t)](K - P_t)\}$                    | $r = 0.06, s = 0.13, K = 45$  | Austin-Brewer <sup>14</sup>          |
| $P_{t+1} = P_t r[K + L \log(P_t) - P_t]$                               | $r = 0.006, K = 600, L = 4$   | Malthus-Condorcet-Mill <sup>15</sup> |

\* Initial conditions are taken uniformly from 0 to 1, except for the Austin-Brewer model, where they are uniform between 0 and 20.

num and the average over all 100 simulations for each model. Because the power varied widely across frequencies, the  $\log_{10}$  of the power is displayed as a function of frequency in Fig. 2.

For each model, the power of fluctuations at high frequencies was at least two orders of magnitude higher, on average, than the power at low frequencies. The maximum power at the lowest frequencies often fell below the minimum power at the highest frequencies. This observation provides assurance that the apparent increase is not due to statistical fluctuations in the spectrum. Although the detailed shape of the average spectrum varied from model to model, and was not always strictly monotonically increasing, all displayed an upward trend in power with increasing frequency. Had a logarithmic scale been used for frequency, as is sometimes the case, the increase in power with increasing frequency would have been still more visually obvious. The greater apparent scatter in the minimum values compared to the maximum values results from the logarithmic transformation of the power.

These results are limited in several respects. A blue spectrum is demonstrated here only for a single point in the parameter space of each model. For most of the models, the parameter values shown are the first values tried that produced chaotic trajectories, suggesting that, within the region of parameter space that produces chaos, many points yield blue spectra. Exploratory numerical calculations suggest that the spectra of chaotic trajectories of these models are blue for many parameter values. However, in the Verhulst model, for example, for a few parameter values the spectrum has a single isolated peak (corresponding to period-3 oscillations) superimposed on a generally blue background (V. Jansen, personal communication), whereas for other parameter values the spectrum is nearly white (G. Sugihara, personal communication). In chaotic trajectories from models other than those considered here, the spectrum may similarly display isolated spikes, depending on parameter values<sup>17</sup>. These results do not cover all nonlinear, iterative, autonomous, deterministic discrete-generation population models that have been or could be considered. However, no models in this class that have been tested so far have had chaotic trajectories with a reddened spectrum, and, for most parameter values, chaotic trajectories that lack discrete peaks are distinctly blue. An analytical understanding of why this is so remains for the future.

The results presented here create a tension between the claims that reddened time series of natural and experimental population sizes are common or ubiquitous, and that chaotic trajectories of nonlinear population (those analysed here and similar) models are accurate descriptions of these time series<sup>18,19</sup>. It is not yet known whether this dilemma can be resolved by expanding the models to take account of environmental fluctuations, the interactions of single species with other species<sup>20</sup>, or the age structure<sup>21,22</sup> and spatial distribution<sup>23</sup> of populations. Periodic and stochastic fluctuations in the parameter  $r$  of the first two models in Table 1 have already been shown to induce complex and counter-intuitive changes in the dynamics of these models<sup>24</sup>, but the spectral effects of such fluctuations in these and additional models have not been investigated. □

Received 14 August; accepted 27 September 1995.

1. Monin, A. S., Kamenkovich, V. M. & Kort, V. G. *Variability of the Oceans* (Wiley, New York, 1977).
2. Jenkins, G. M. & Watts, D. G. *Spectral Analysis and its Applications* (Wiley, New York, 1968).
3. Steele, J. H. *Nature* **313**, 355–358 (1985).
4. Steele, J. H. *J. theor. Biol.* **153**, 425–436 (1991).
5. Steele, J. H. & Henderson, E. W. *Science* **224**, 985–987 (1984).
6. Steele, J. H. & Henderson, E. W. *Phil. Trans. R. Soc. Lond. B* **343**, 5–9 (1994).
7. Pimm, S. L. & Redfearn, A. *Nature* **334**, 613–614 (1988).
8. Arino, A. & Pimm, S. L. *Evol. Ecol.* **9**, 429–443 (1995).
9. Caswell, H. & Cohen, J. E. *J. theor. Biol.* **176**, 301–316 (1995).
10. Halley, J. M. *Trends Ecol. Evol.* (in the press).
11. Lewontin, R. C. & Cohen, D. *Proc. natn. Acad. Sci. U.S.A.* **62**, 1056–1060 (1969).
12. May, R. M. & Oster, G. *Am. Nat.* **110**, 573–599 (1976).
13. Hassell, M. P., Lawton, J. H. & May, R. M. *J. Anim. Ecol.* **45**, 471–486 (1976).
14. Austin, A. L. & Brewer, J. W. *Technol. Forecast. soc. Change* **3**, 23–49 (1971).
15. Cohen, J. E. *Science* **269**, 341–348 (1995).
16. Li, T.-Y., Misiurewicz, M., Pianigiani, G. & Yorke, J. A. *Phys. Lett. A* **87**, 271–273 (1982).
17. Farmer, D., Crutchfield, J., Froehling, H., Packard, N. & Shaw, R. *Ann. N.Y. Acad. Sci.* **357**, 453–472 (1980).
18. Schaffer, W. M. in *Chaos in Biological Systems* (eds Degn, H., Holden, A. V. & Olsen, L. F.) 233–248 (NATO Advanced Science Institutes Series A, vol. 138) (Plenum, New York, 1987).
19. Hastings, A., Hom, C. L., Ellner, S., Turchin, P. & Godfray, H. C. J. *A. Rev. Ecol. Syst.* **24**, 1–33 (1993).
20. Muratori, S. & Rinaldi, S. *SIAM J. appl. Math.* **52**, 1688–1706 (1992).
21. Levin, S. A. & Goodyear, C. P. *J. math. Biol.* **9**, 245–274 (1980).
22. Gurney, W. S. C., Nisbet, R. M. & Lawton, J. H. *J. Anim. Ecol.* **52**, 479–495 (1983).
23. Ruxton, G. D., Bascompte, J. & Solé, R. V. *J. Anim. Ecol.* **63**, 1002–1003 (1994).
24. Markus, M., Hess, B., Rössler, J. & Kiwi, M. in *Chaos in Biological Systems* (eds Degn, H., Holden, A. V. & Olsen, L. F.) 267–277 (NATO Advanced Science Institutes Series A, vol. 138) (Plenum, New York, 1987).
25. Moran, P. A. P. *Biometrics* **6**, 250–258 (1950).
26. Ricker, W. E. *J. Fish. Res. Bd Can.* **11**, 559–623 (1954).
27. Pennycuik, C. J., Compton, R. M. & Beckingham, L. J. *J. theor. Biol.* **18**, 316–329 (1968).
28. Hassell, M. P. *J. Anim. Ecol.* **44**, 283–296 (1974).
29. Maynard Smith, J. *Models in Ecology* (Cambridge Univ. Press, 1974).
30. Varley, G. C., Gradwell, G. R. & Hassell, M. P. *Insect Population Ecology* (Blackwell, Oxford, 1973).

ACKNOWLEDGEMENTS. I thank A. E. Cohen, J. M. Halley, V. A. Jansen, J. H. Lawton, R. M. May, G. Sugihara and an anonymous referee for help during this work, and W. T. Golden and S. L. Golden for hospitality. This work was supported by the U.S. National Science Foundation and the NERC Centre for Population Biology, Imperial College.

## Visually evoked calcium action potentials in cat striate cortex

Judith A. Hirsch, José-Manuel Alonso & R. Clay Reid

The Rockefeller University, Laboratory of Neurobiology, 1230 York Avenue, New York, New York 10021, USA

EARLY intracellular studies of cerebral cortical neurons indicated that synaptic input evokes dendritic action potentials that convey information towards the soma<sup>1</sup>. Subsequent work *in vitro* established that neocortical neurons produce dendritic  $\text{Ca}^{2+}$  action potentials<sup>2–5</sup>. To determine whether natural stimuli elicit  $\text{Ca}^{2+}$  spikes, we combined the techniques of whole-cell recording<sup>6,7</sup>, pharmacology<sup>8</sup> and quantitative receptive field mapping<sup>9</sup>. Our findings show that visual stimulation routinely evoked  $\text{Ca}^{2+}$  spikes in distinct functional<sup>10</sup> and anatomical<sup>11</sup> classes of cells in different layers of the cat striate cortex<sup>12</sup>. Hence regenerative  $\text{Ca}^{2+}$  potentials appear to play a role in both the initial and later stages of cortical sensory processing.

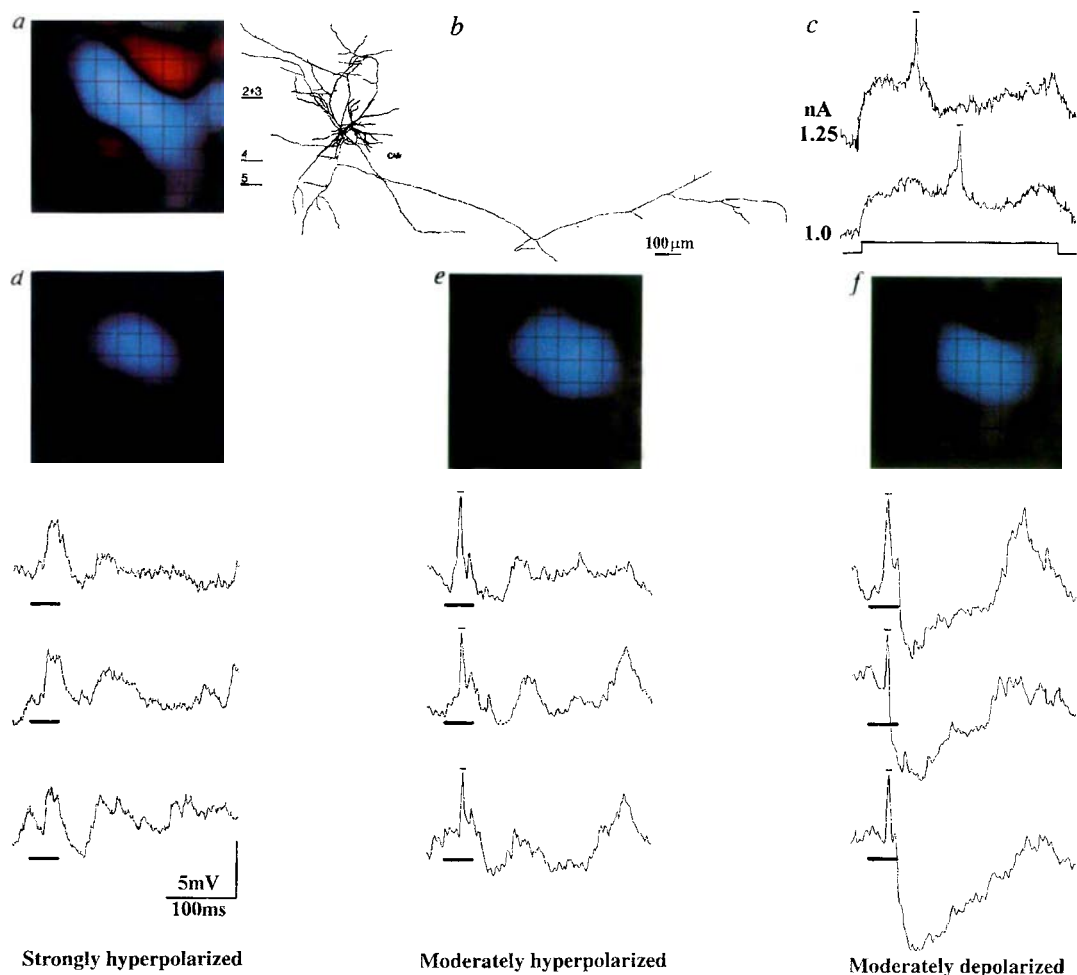


FIG. 1 A simple cell studied during intracellular blockade of the Na<sup>+</sup> spike. *a*, The receptive field had an off centre (blue) and two on flanks (red). The black lines indicate the grid spacing, which was 0.85°; the data were smoothed by 0.5 pixel. Each square was 1.7° wide and 39 ms in duration. *b*, The cell was a spiny stellate neuron whose dendrites ramified mainly in layer 4. Vertical projections were dense in the superficial layers and sparse in layer 6; horizontal connections terminated in layer 6. *c*, Responses to current injection. Thin bars above the spikes mark the portion of the trace selected by the template. *d–f*, Top panels map the portions of the receptive field where the visual stimulus evoked Ca<sup>2+</sup> spikes when the membrane was polarized: *d*, -1.0 nA; *e*, -0.5 nA; *f*, +0.4 nA. All responses illustrated in the bottom panels of *d–f* were evoked by a dark square presented to the same site in the middle right of the off subfield. Each panel shows three trials of the stimulus, marked by the thick lines: *d*, from a stimulus run made 64–69 min into the recording; *e* was made 8–13 min and *f*, 110–115 min into the recording. The average latency to the peak region of the spike for the pixel shown was  $21.0 \pm 2$  ms. Ca<sup>2+</sup> action potentials were recorded in two other spiny stellate cells; the input resistance  $R_i$  was  $85 \pm 64$  MΩ, and  $\tau$  (time constant) was  $24 \pm 7$  ms.

**METHODS.** Adult cats were anaesthetized with intramuscular ketamine (10 mg kg<sup>-1</sup>) and intravenous sodium pentothal (20 mg kg<sup>-1</sup>; supplemented at rate of 2 mg kg<sup>-1</sup> h<sup>-1</sup>) and were paralysed with vecuronium

bromide (0.2 mg kg<sup>-1</sup> h<sup>-1</sup>). Each stimulus run consisted of light and dark squares at 100% contrast flashed singly for 29–39 ms in pseudo-random order 16 times on a 16 × 16 grid<sup>9</sup>. Grid spacing ranged from 0.4–0.85° and square size from 0.4–1.7°. Maps of the entire receptive field were calculated from averaged intracellular records. Maps of the Ca<sup>2+</sup> action potentials were made from events labelled by simple templates designed to distinguish the spikes from other events. The templates had three parameters determined from the shape of representative spikes: peak, decline and width. A template match was positive if the voltage first increased by the peak value and then decreased by the decline (smaller than the peak) all within the width (7–10 ms; variation is among cells). In order to map the spatial distribution of the spikes, the timing of the event was defined as the beginning of the window. Pipette resistance was  $\geq 12$  MΩ when filled with internal solution (concentrations in mM): K<sup>+</sup> gluconate, 120; NaCl, 5; CaCl<sub>2</sub>, 1; MgCl<sub>2</sub>, 1; EGTA, 11; GTP, 0.2; ATP, 2; HEPES, 40; QX-314 Br (courtesy of Astra), 10; biocytin 1%, pH 7.3, 290 mOsm. Recordings were made with an Axopatch 200a amplifier and stored digitally; neither capacitance nor series resistance was compensated for, so fast events were filtered and the bridge was balanced off line. After processing for histology, labelled neurons were drawn using a camera lucida. Reconstruction of the electrode tracks revealed that, with one exception where the proximal apical dendrite was patched, our recordings came from the soma.

Long-lasting intracellular recordings with patch electrodes gave sufficient time to map receptive fields precisely and to monitor changes in response under varied experimental conditions (0.4–2.5 h;  $n=19$ ). It was also possible to label the neurons under study and resolve their morphology and laminar position. To determine whether regenerative Ca<sup>2+</sup> potentials were triggered by visual stimulation, we recorded from neurons in which Na<sup>+</sup> conductances were blocked by intracellular

perfusion with the lidocaine derivative QX-314 (ref. 8). Figure 1 shows results obtained from a spiny stellate cell (Fig. 1*b*) studied during Na<sup>+</sup> channel blockade; recordings were made from the soma over a 2-hour interval. The neuron was in layer 4, which receives the densest projection from the lateral geniculate nucleus of the thalamus. It had a simple receptive field<sup>10</sup> (Fig. 1*a*) with an off centre (blue) and two weaker on flanks (red). The receptive field was mapped by plotting the

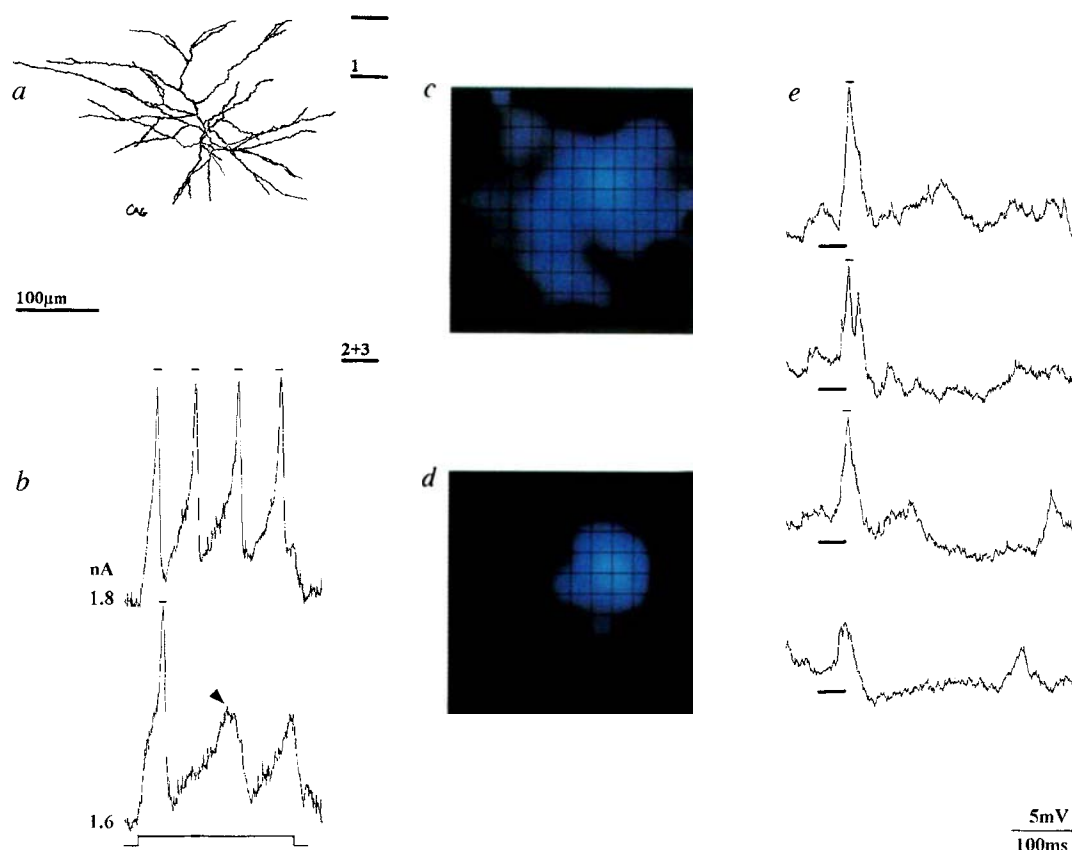


FIG. 2 A complex cell recorded during intracellular blockade of the  $\text{Na}^+$  spike. *a*, The cell was a pyramid whose dendritic arbor extended through the upper portion of layer 2+3 and layer 1; the axon was not fully reconstructed. *b*, Responses to injection of depolarizing pulses of current. Arrowhead denotes a subliminal oscillation. Thin bars above the spikes mark the portion of the trace selected by the template. *c*, The receptive field was roughly circular; the cell responded only to dark stimuli. The grid spacing was  $0.85^\circ$ ; each square was  $1.7^\circ$  wide and 39 ms in duration. *d*, Spatial distribution of the  $\text{Ca}^{2+}$  spikes elicited during the same stimulus run. *e*, Four trials of a stimulus presented to

the centre of the receptive field; the last trace shows an episode that remained subthreshold for the spike. Thick bars mark the stimulus. The average latency to the peak region of the spike for the pixel illustrated was  $41.3 \pm 4.8$  ms.  $\text{Ca}^{2+}$  action potentials have been recorded in two other superficial pyramidal cells and three other morphologically unidentified complex cells;  $R_i = 84 \pm 74 \text{ M}\Omega$ ,  $\tau = 21 \pm 6$  ms. Responses to current injection were obtained in the first 3–5 min of the recording. Visually evoked responses were collected during an interval lasting ~10–15 min after the recording began.

averaged changes in membrane potential evoked by visual stimulation.

Depolarizing current pulses evoked small, broad, QX-314-resistant events, presumed to be  $\text{Ca}^{2+}$  action potentials (Fig. 1c). Visual stimuli also triggered these spikes (Fig. 1d–f). Using a template fitted to the action potentials that were evoked electrically, we labelled events that were driven visually and mapped their spatial distribution within the receptive field (Fig. 1d–f, top). To emphasize the voltage dependence of the action potentials, these maps were made when the cell was injected with different amounts of current (Fig. 1d–f, top). Strong hyperpolarization inhibited all but a few spikes evoked by stimuli that fell in the peak of the receptive field (Fig. 1d, top). Thus, the plot of the spikes is much smaller than that of the entire receptive field, which includes subthreshold postsynaptic potentials (panels *a* and *d* from Fig. 1 were made from the same stimulus run). Figure 1d (bottom) shows three individual responses to a dark square presented immediately to the right of the field's peak. These traces reveal the graded shapes of subthreshold responses. During moderate hyperpolarization, visually evoked spikes were produced throughout most of the *off* subregion (Fig. 1e, top); a second action potential seemed to follow the initial one, but was broader than the template (Fig. 1e, bottom). When the membrane was depolarized,  $\text{Ca}^{2+}$  spikes were produced both

in the *off* subregion and in the stronger of two *on* flanks (Fig. 1f, top).

The early appearance of the visually evoked spikes suggested that they were generated by geniculocortical inputs, which convey ascending information to cortex. Although we could not identify the subcellular origins of the calcium spikes that we recorded, work *in vitro* (refs 2–5, 13 for example) associates them with the dendrites. Also, we often detected multiple sets of  $\text{Ca}^{2+}$  spikes<sup>3,5</sup> in visually evoked responses (for example, Fig. 1e) but usually produced just one type with depolarizing current pulses (for example, see Fig. 1c). Presumably, direct current injection influenced only the trigger site nearest the electrode.

Visual responses of complex cells whose inputs appeared to be mainly, if not exclusively, intracortical in origin were also studied during blockade of the  $\text{Na}^+$  spike. Results obtained from one neuron are illustrated in Fig. 2. The dendritic arbor was confined to the top of layer 2+3 (Fig. 2a), out of the range of the fibres from the primary layers of the lateral geniculate; the recording site was at the soma. The receptive field obtained from the averaged intracellular voltage is shown in Fig. 2c. The distribution of the evoked  $\text{Ca}^{2+}$  spikes (Fig. 2d) was smaller, restricted to the peak of the field. All four traces in Fig. 2e were collected during the same stimulus run. The top three traces show instances where the stimulus elicited an action potential. The bottom trace illustrates a subthreshold response.



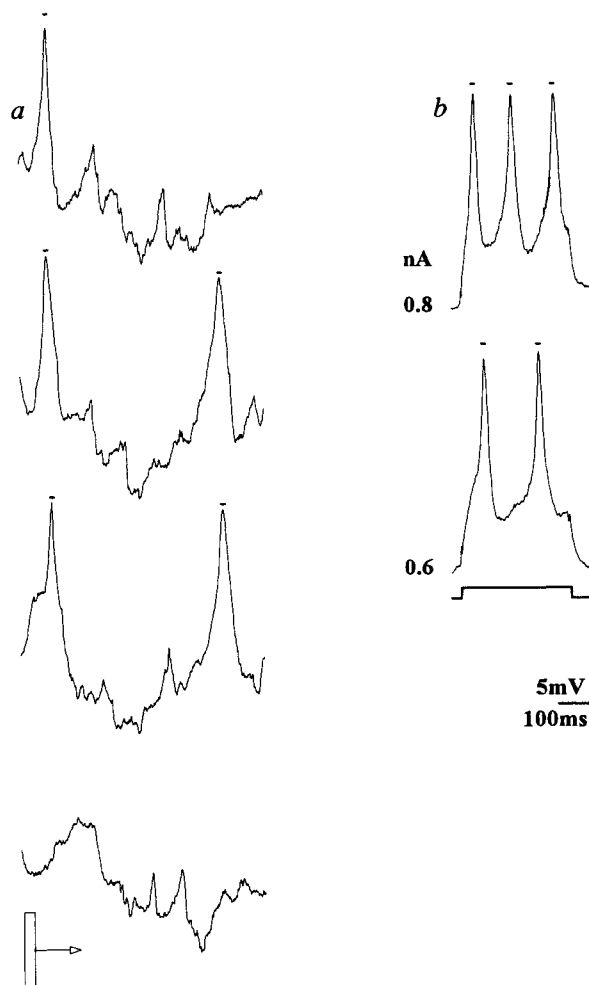


FIG. 3  $\text{Ca}^{2+}$  action potentials evoked by moving stimuli during intracellular  $\text{Na}^+$  blockade. *a*, Responses of a complex cell to a light  $1.70^\circ$  bar swept through the receptive field at  $11.1^\circ \text{ s}^{-1}$ . The top three traces show stimulus-evoked action potentials and the bottom trace depicts a subthreshold response. Thin bars above the spikes mark the portion of the trace selected by the template. *b*, Responses to depolarizing pulses of current. Responses to current injection were obtained in the initial 3–5 min of the recording. Visually evoked responses were collected ~14–18 min after the recording began.

Some complex cells were poorly excited by flashed squares but responded vigorously to moving oriented stimuli. The behaviour of one such cell to a bright bar of the optimal orientation swept across the receptive field is shown in Fig. 3*a*. The bar evoked a small broad action potential as it entered and sometimes as it left the receptive field. Although this cell was not morphologically identified, the electrode track showed that the recording was made in superficial layer 2 + 3.

Recordings made in the absence of  $\text{Na}^+$  channel blockade indicated that dendritic action potentials are normally evoked during sensory processing. The records shown in Fig. 4 were obtained from the soma of an individual simple cell. A light stimulus presented within the *on* subfield (Fig. 4*b*) evoked large, fast,  $\text{Na}^+$  spikes (Fig. 4*c*). Close inspection revealed that these conventional action potentials were often preceded by shorter waveforms (Fig. 4*c*, arrows) that were reminiscent of 'prepotentials', putative dendritic spikes observed in earlier studies<sup>1,14,15</sup>. Spontaneous activity, however, showed only fast spikes and excitatory postsynaptic potentials (Fig. 4*d*). The quality of this recording declined slightly after ~40 min, and small broad action potentials that resembled the prepotentials appeared during spontaneous activity (Fig. 4*e*, arrows; compare refs 14, 16, 17). The slowness of the small spikes recalled dendritic  $\text{Ca}^{2+}$  potentials identified *in vitro*<sup>13,18,19</sup>. Hyperpolarizing the cell made it possible to view the small spikes in isolation and to monitor the time course of their appearance following visual stimulation (Fig. 4*f*). Their presence at the start of the response is consistent with the interpretation that they were generated by the thalamic afferents. Only the short spikes persisted during somatic hyperpolarization, supporting the idea that the two classes of action

potentials arose at separate places. The small spikes were again visible as prepotentials after the hyperpolarizing current was relieved (Fig. 4*g*).

When  $\text{Na}^+$  channels were silenced with QX-314, we found that visual stimulation evoked  $\text{Ca}^{2+}$  action potentials in striate cortical neurons. *In vitro*, the  $\text{Ca}^{2+}$  spike has been found to have a higher threshold than either the dendritic or axonal  $\text{Na}^+$  spike<sup>4</sup> when current is injected into the dendrite. Hence it has been proposed that action potentials that originate at sites removed from the initial segment serve as retrograde signals to indicate that the axon has fired<sup>4</sup>. Although this must often be the case, it is important to remember that distributed synaptic inputs compartmentalize the cortical dendritic arbor<sup>20–22</sup>. Thus the calcium spike, or dendritic  $\text{Ca}^{2+}$  and  $\text{Na}^+$  potentials in various balance<sup>5,23,24</sup>, may be particularly useful in traversing regions of the membrane where input resistance is reduced by vigorous synaptic input. This view fits well with our observation in control recordings that prepotentials normally occurred during strong visual drive (Fig. 4), when the membrane was probably perforated by opened transmitter-gated channels.

The finding that synaptic input activates regenerative  $\text{Ca}^{2+}$  potentials is important in considering the impact of a given projection. For example, geniculocortical inputs may contribute as few as 5% of the excitatory synapses made in layer 4 (refs 25, 26). Nonetheless, cross-correlation studies indicate that the influence of a single primary afferent on its target simple cells is strong<sup>27,28</sup>. Part of this strength may arise from interactions between thalamocortical excitation and the regenerative conductances it engages. Additionally, local changes in calcium may set in motion diverse metabolic processes that modify synaptic

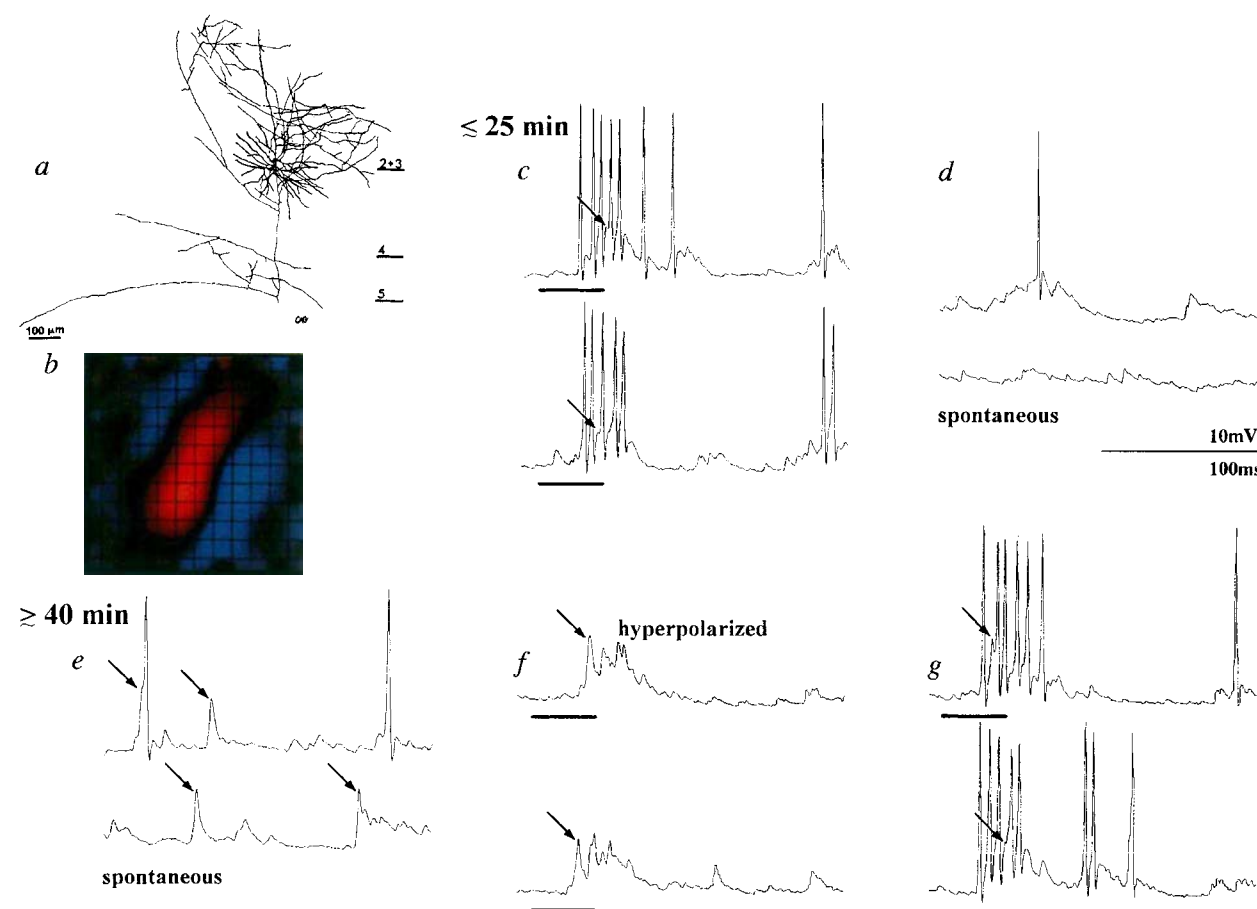


FIG. 4 A simple cell that produced multiple types of action potential in the absence of  $\text{Na}^+$  channel blockade. *a*, The cell was a pyramid whose soma lay at the border between layers 2 + 3 and 4. Its apical dendrite extended through the superficial layers and its basal dendrites arborized in layer 4. The bulk of the axon projected superficially while sparser horizontal processes coursed through layers 2 + 3 and 5. *b*, The receptive field had an on centre with two off flanks. The grid spacing was  $0.4^\circ$ ; each square was  $0.8^\circ$  across and 29 ms in duration. The responses seen in *c*, *f* and *g* were evoked by a light stimulus shone at in the centre of the on subregion. Each trace shows an individual response to different trials of the same stimulus; thick bars mark the stimulus. *c*, Responses obtained within the first 25 min of recording. Arrows indicate examples of the short waveforms that sometimes

appeared to precede the larger spikes. *d*, Spontaneous activity during the first half of the recording. *e*, Spontaneous activity observed later in the recording included two types of action potentials; large fast ones and shorter slower ones (arrows) that resemble the prepotentials seen in *c*. *f*, The evoked response contained only small spikes when the membrane was hyperpolarized with  $-0.5 \text{ nA d.c.}$ . *g*, Relief of the current restored the larger impulses; again the arrows mark the shorter spike. Visually evoked prepotentials were seen in three other simple cells ( $R_1 = 88 \pm 45 \text{ M}\Omega$ ,  $\tau = 23 \pm 5 \text{ ms}$ ), one spiny stellate cell in layer 4 and two pyramids in upper layer 6; and in five complex pyramidal cells ( $R_1 = 69 \pm 27 \text{ M}\Omega$ ;  $\tau = 16 \pm 4 \text{ ms}$ ) in layers 2 + 3, 5 and 6. Methods were the same as for Fig. 1 except that QX-314 Br was omitted from the internal solution.

function<sup>29</sup>. As we have recorded calcium action potentials from simple cells as well as from superficial complex cells, we believe that these spikes provide a means of influencing both ascending and intracortical input. □

Received 6 October; accepted 16 October 1995.

- Spencer, W. A. & Kandel, E. R. *J. Neurophysiol.* **24**, 272–285 (1961).
- Stafstrom, C. E., Schwindt, P. C., Chubb, M. C. & Crill, W. E. *J. Neurophysiol.* **53**, 153–170 (1985).
- Reuveni, I., Friedman, A., Amitai, Y. & Gutnick, M. J. *J. Neurosci.* **13**, 4609–4621 (1993).
- Stuart, G. & Sakmann, B. *Nature* **367**, 69–72 (1994).
- Yuste, R., Gutnick, M., Saar, D., Delaney, K. & Tank, D. *Neuron* **13**, 23–43 (1994).
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Archiv ges. Physiol.* **391**, 85–100 (1981).
- Ferster, D. & Jagadeesh, B. *J. Neurosci.* **14**, 1262–1274 (1992).
- Strichartz, G. R. *J. gen. Physiol.* **62**, 37–57 (1973).
- Jones, J. P. & Palmer, L. A. *J. Neurophysiol.* **58**, 1187–1211 (1987).
- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **160**, 106–154 (1962).
- Lorente De Nö, R. in *Cerebral Cortex* (ed. Fulton, J. F.) 291–325 (Oxford University Press, London, 1944).
- Hirsch, J. A., Alonso, J. M. & Reid, R. *Soc. Neurosci. Abstr.* **20**, 1476 (1994).
- Amitai, Y., Friedman, A., Connors, B. W. & Gutnick, M. J. *Cerebral Cortex* **3**, 26–38 (1993).
- Purpura, D. & Shofar, R. *J. Neurophysiol.* **27**, 117–132 (1963).

- Llinas, R. & Nicholson, C. *J. Neurophysiol.* **34**, 532–551 (1971).
- Pockberger, H. *Brain Res.* **539**, 181–190 (1991).
- Steriade, M., Nunez, A. & Amzica, F. *J. Neurosci.* **13**, 3252–3265 (1993).
- Llinas, R. & Sugimori, M. *J. Physiol., Lond.* **305**, 197–213 (1980).
- Regehr, W., Kehoe, J., Ascher, P. & Armstrong, C. *Neuron* **11**, 145–151 (1993).
- Bernander, O., Douglas, R. J., Martin, K. M. & Kock, C. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11569–11573 (1991).
- Traub, R., Jefferys, J. R., Miles, R., Whittington, M. & Toth, K. *J. Physiol., Lond.* **481**, 79–95 (1994).
- Borst, A. & Engelhaaf, M. *Proc. natn. Acad. Sci. U.S.A.* **89**, 4139–4143 (1992).
- Jaffe, D. B., Lasser-Ross, N., Lisman, J. E., Miyakawa, H. & Ross, W. N. *Nature* **357**, 244–246 (1992).
- Regehr, W. & Tank, D. W. *J. Neurosci.* **12**, 4202–4223 (1992).
- Peters, A. & Payne, B. R. *Cerebral Cortex* **3**, 69–78 (1993).
- Ahmed, B., Anderson, J. C., Douglas, R. J., Martin, K. M. & Nelson, J. C. *J. comp. Neurol.* **341**, 39–49 (1994).
- Tanaka, K. *J. Neurophysiol.* **49**, 1303–1318 (1983).
- Reid, R. C. & Alonso, J. M. *Nature* **378**, 281–284 (1995).
- Ghosh, A. & Greenberg, M. *Science* **268**, 239–247 (1995).
- Hirsch, J. J. *J. Physiol., Lond.* **483**, 183–199 (1995).

ACKNOWLEDGEMENTS. We thank T. N. Wiesel for contribution to all phases of the project; M. Nowicky for suggestions to improve the manuscript, C. Gallagher for drawing labelled cells, P. Peirce for photographing them, and C. Gallagher and K. McGowan for help with histology. Supported by the NIH, the Klingenstein Fund Inc., Fulbright/MEC, and the Charles H. Revson Foundation.

# Impairment of antigen-specific T-cell priming in mice lacking CD40 ligand

Iqbal S. Grewal, Jianchao Xu & Richard A. Flavell

Howard Hughes Medical Institute and Section of Immunobiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA

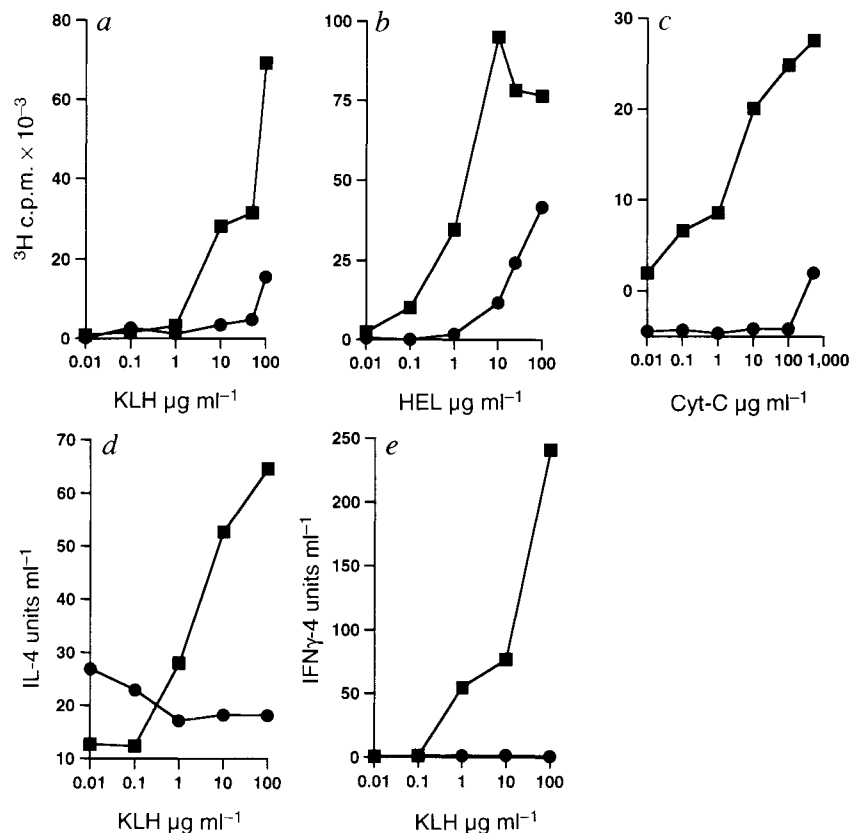
**LACK of functional expression of CD40 ligand (CD40L) on T cells results in hyper-IgM syndrome (HIGM1), a human immunodeficiency associated with a severely impaired humoral immune response that is consistent with defects in B-cell responses<sup>1-3</sup>. Patients also succumb to recurrent opportunistic infections such as *Pneumocystis carinii* and *Cryptosporidial* diarrhoea<sup>4,5</sup>, suggesting that T-cell functions are also compromised in these individuals, but so far this has not been explained. We have previously shown that mice deficient for CD40L, like HIGM1 patients, show grossly abnormal humoral responses<sup>6</sup>. Here we report that CD40L-deficient mice are defective in antigen-specific T-cell responses. Adoptively transferred antigen-specific CD4<sup>+</sup> T cells lacking CD40L failed to expand upon antigen challenge of the recipients, showing that expression of CD40L on T cells is required for *in vivo* priming of CD4<sup>+</sup> T cells and therefore for the initiation of specific T-cell immune responses.**

To determine whether CD40L influences T-cell responsiveness, we immunized CD40L-deficient and wild-type mice with a protein antigen, keyhole limpet haemocyanin (KLH), and tested their *in vitro* recall proliferative response. A significantly reduced response was observed in CD40L-deficient mice compared with wild-type mice (Fig. 1a) and the antigen dose-response curve was shifted by several orders of magnitude. The same defect was demonstrated with two other protein antigens, hen egg lysozyme (HEL) and cytochrome *c* (Cyt-*c*) (Fig. 1b, c).

Furthermore, a dramatic reduction in the production of both interleukin-4 (IL-4) and interferon- $\gamma$  (IFN- $\gamma$ ) was seen in the CD40L-deficient mice (Fig. 1d, e), indicating a key role for CD40L in antigen-specific T-cell responses.

The reduced response in CD40L-deficient mice could either be caused by a problem with the antigen-presenting cells (APC) or with T cells. We therefore used APC from wild-type mice to activate CD4<sup>+</sup>-enriched T cells from CD40L-deficient mice immunized with KLH/CFA (Fig. 2a, b). Neither wild-type APC (Fig. 2a) nor lipopolysaccharide (LPS)-activated B cells (which express co-stimulatory activity and thus might bypass a CD40L requirement) were able to restore the defects (Fig. 2c). To address whether APC are intrinsically defective in CD40L-deficient mice, we adoptively transferred wild-type CD4<sup>+</sup> T cells to CD40L-deficient mice; these transferred wild-type CD4<sup>+</sup> T cells responded vigorously to KLH immunization in the CD40L-deficient recipient (Fig. 2d). These observations suggest that the *in vivo* defect lies at the level of the T cells rather than APC. Despite this defect, CD40L-deficient mice do make a weak proliferative response to antigens, suggesting that a CD40L-independent mechanism for T-cell activation must also exist *in vivo*.

To determine whether CD40L-deficient T cells have an intrinsic defect in the ability to proliferate, responses of naive T cells to polyclonal activators were analysed. Naive CD4<sup>+</sup> T cells from both wild-type and CD40L-deficient mice proliferated normally in response to anti-CD3, concanavalin A (conA), SEA or SEB (Fig. 3a-d). To test the *in vitro* primary response to antigen, we crossed CD40L-deficient mice to a Cyt-*c*-specific-TCR transgenic mice, and obtained mice with T cells specific for Cyt-*c* but lacking CD40L. Proliferation of Cyt-*c*-specific CD4<sup>+</sup> TCR transgenic T cells from wild-type and from CD40L-deficient mice was identical (Fig. 3e). Thus, the intrinsic potential of T cells to respond to antigens is preserved in CD40L-deficient mice. The number and ratios of various T-cell subsets in CD40L-deficient mice are normal, as shown earlier<sup>6</sup>. In addition, the ability to obtain functional Cyt-*c*-specific T cells from



**FIG. 1** CD40L mice have impaired lymph node recall proliferative responses to protein antigens. *In vitro* proliferative recall responses of wild-type (■) and CD40L-deficient (●) mice to KLH (a), to HEL (b), or to Cyt-*c* (c). Production of IL-4 (d) and IFN- $\gamma$  (e) by purified CD4<sup>+</sup> cells from draining lymph nodes of wild-type and CD40L-deficient mice immunized with KLH/CFA to *in vitro* challenge of KLH.

**METHODS.** Mice were immunized with 100  $\mu\text{g}$  KLH, HEL or Cyt-*c* in complete Freund's adjuvant (CFA) in the hind footpads, and 9 days later draining lymph nodes were tested for recall proliferative responses as described<sup>14</sup>. CD4<sup>+</sup> T cells were purified from draining lymph nodes as described<sup>15</sup>. Assays for cytokine production by T cells from KLH-immunized mice were conducted by culturing purified CD4<sup>+</sup> T cells with APC in the presence of the indicated amount of KLH. After 4 days, supernatant IL-4 and IFN- $\gamma$  was assayed using an ELISA kit as recommended by the manufacturer (Pharmingen). Experiments were repeated at least three times.

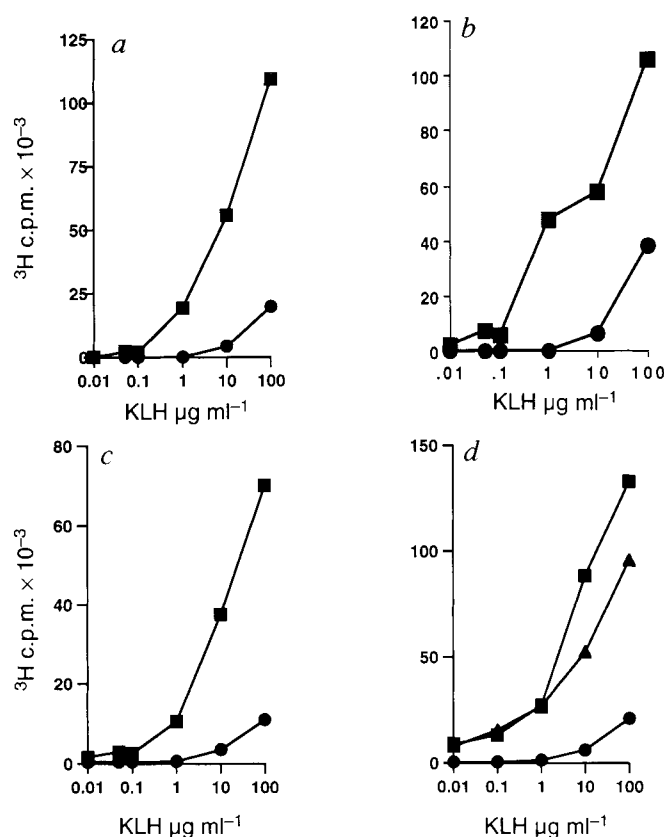
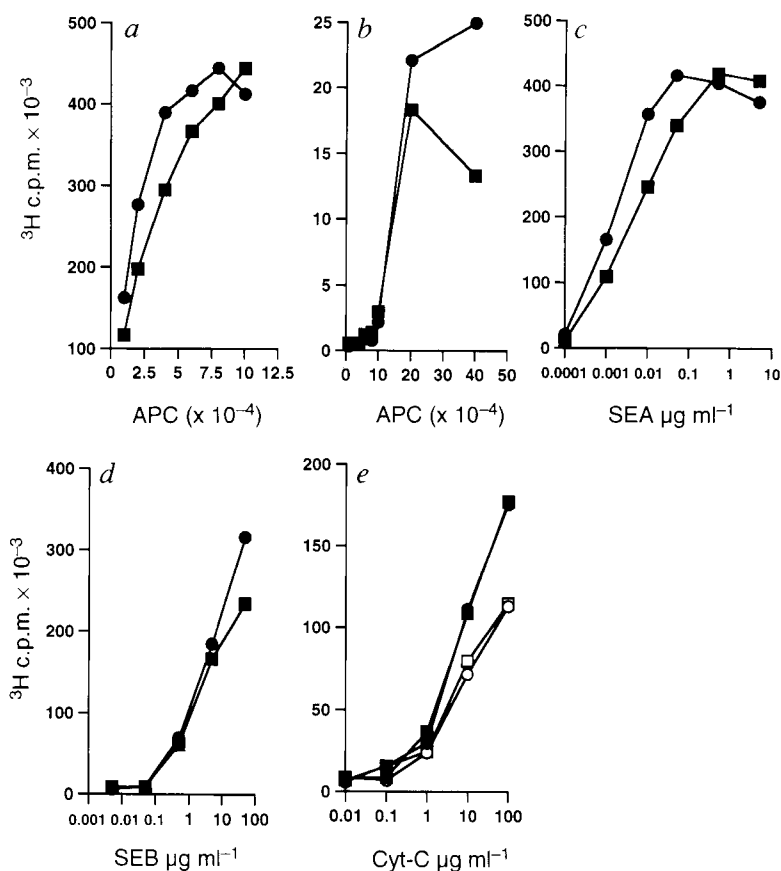


FIG. 2 Lack of response of CD40L-deficient mice to antigen immunization is due to a T-cell defect. Responses are shown of purified CD4<sup>+</sup> T cells from KLH/CFA-immunized wild-type mice (■) or CD40L-deficient mice (●) to KLH in the presence of control APC (a) or APC from CD40L-deficient mice (b); responses are also shown of purified CD4<sup>+</sup> T cells to KLH in the presence of LPS-activated splenic B cells as APC (c). d, Responses to KLH of CD40L-deficient mice (▲) that received  $1 \times 10^7$  CD4<sup>+</sup> T cells i.v. from wild-type mice, and of wild-type and CD40L-deficient mice.

METHODS. Proliferation was measured by culturing  $1 \times 10^5$  CD4<sup>+</sup> T cells from KLH-immunized mice with  $5 \times 10^5$  irradiated (3,000 R) spleen cells from unimmunized wild-type or from CD40L-deficient mice or with LPS-activated ( $50 \mu\text{g ml}^{-1}$  for 48 h) B cells from wild-type mice in the presence of indicated amounts of KLH. Proliferation was determined after 4 d of culture by incorporation of <sup>3</sup>H-thymidine.

FIG. 3 Naive T cells from CD40L-deficient mice respond normally to antigenic and polyclonal stimuli. Lymph-node T cells purified from unimmunized wild-type (■) or CD40L-deficient (●) mice were cultured in the presence of a, conA; b, anti-CD3; c, SEA; or d, SEB. e, Lymph node CD4<sup>+</sup> T cells purified from Cyt-c-specific TCR transgenic mice lacking CD40L (circles) or control mice (squares) were tested for *in vitro* proliferative responses to Cyt-c in the presence of APC from wild-type mice (filled symbols), and from CD40L-deficient mice (open symbols).

METHODS. Naive lymph node cells were cultured in the presence of wild-type APC at different concentrations in the presence of conA ( $3 \mu\text{g ml}^{-1}$ ) or anti-CD3 ( $1 \mu\text{g ml}^{-1}$ ), and in the presence of various concentrations of SEA or SEB. Proliferation was determined after 4 days of culture by incorporation of <sup>3</sup>H-thymidine. For testing naive T cells from TCR transgenic mice, female mice homozygous for the CD40L mutation were crossed with Cyt-c-specific-TCR transgenic mice and progeny were screened for the presence of Cyt-c-specific TCR by staining with Vβ3 and Vα11 antibodies. Cyt-c-specific T cells from either CD40L-deficient or from wild-type control mice were cultured in the presence of APC from wild-type or CD40L-deficient mice together with different concentrations of Cyt-c. Proliferation was determined after 4 days of culture by incorporation of <sup>3</sup>H-thymidine.





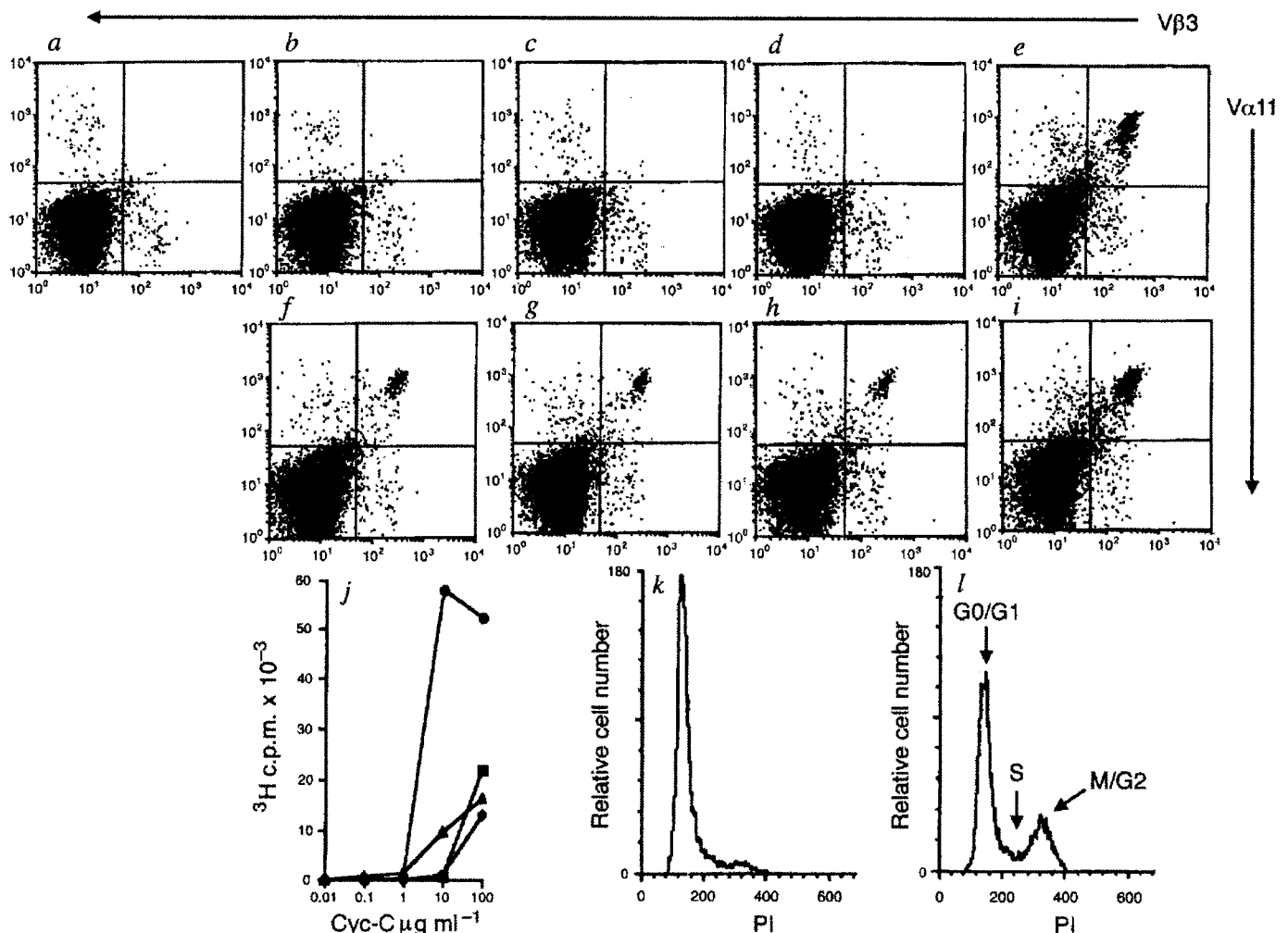


FIG. 4 Adoptive transfer of wild-type or CD40L-deficient Cyt-c-specific-TCR transgenic cells to recipient mice. FACS analysis of draining lymph node cells from mice that did not receive TCR transgenic T cells but were immunized with Cyt-c in CFA (a), and from mice that received  $5 \times 10^6$  CD40L-deficient (b, d) or wild-type (c, e) TCR transgenic lymph node cells, and from mice that received  $1.1 \times 10^7$  CD40L-deficient (f, h) or wild-type (g, i) TCR transgenic lymph node cells. Recipient mice were also immunized with Cyt-c in CFA (d, e, h, i). *In vitro* recall responses of draining lymph nodes to Cyt-c (j) from mice that received  $1.1 \times 10^7$  CD40L-deficient (◆) or wild-type (■) TCR transgenic lymph node cells, and from mice that were immunized with Cyt-c in CFA and also received  $1.1 \times 10^7$  CD40L-deficient (▲) or wild-type (●) TCR transgenic lymph node cells. k, Flow cytometry plots of cell-cycle analysis for draining lymph node cells by propidium iodide staining for mice that

received  $1.1 \times 10^7$  TCR transgenic lymph node cells lacking CD40L and were immunized with Cyt-c/CFA, and l, mice that received  $1 \times 10^7$  wild-type TCR<sup>+</sup> T cells and were immunized with Cyt-c/CFA. METHODS. Cyt-c TCR transgenic lymph node cells from mice lacking CD40L or from control mice were adoptively transferred i.v. into recipient mice (TCR-negative littermates). A set of recipient mice were also immunized with Cyt-c in CFA (100 μg per mouse in the hind footpads). Five days after transfer, draining lymph node cells were stained with Vβ3 and Vα11 antibodies. Recipient mice that received TCR transgenic lymph nodes but were not immunized with Cyt-c, and mice that did not receive TCR transgenic cells but were immunized with Cyt-c were also tested as controls in the experiments. *In vitro* proliferation of draining lymph node cells was also measured in the presence of Cyt-c. Cell cycle monitoring was as ref. 16.

TCR transgenic CD40L-deficient mice indicates that positive selection of transgenic TCR-specific T cells does not require CD40L.

The failure of the apparently functional T cells from CD40L-deficient mice to respond to antigen stimulation *in vivo* might be explained by an inefficient *in vivo* priming of antigen-specific T cells in CD40L-deficient mice. We therefore adoptively transferred a limiting number of T cells from Cyt-c-specific-TCR transgenic mice into recipients immunized with Cyt-c (ref. 7) and examined their draining lymph nodes for the presence of transgenic TCR-specific T cells (Fig. 4a-i) and for Cyt-c-specific T-cell responses (Fig. 4j) five days after transfer. T cells lacking CD40L failed to expand *in vivo* upon antigen challenge, whereas wild-type cells expanded normally (Fig. 4a-i). The same results were seen when expansion was monitored 3 days postimmunization (data not shown). To test whether the numbers of T cells

from CD40L-deficient mice had decreased after adoptive transfer (for example as a result of apoptosis or a graft-rejection artefact), we repeated the experiments using a higher dose of input T cells. Again, expansion was seen for wild-type CD4<sup>+</sup> T cells but not for the CD40L-deficient T cells. The initial population of transferred cells does not, however, diminish, suggesting that CD40L-deficient T cells fail to expand but do not die. Consistent with this, the *in vitro* T-cell response of CD40L-deficient T cells recovered from draining lymph nodes, which is not defective (Fig. 3e), can easily be seen for the adoptively transferred CD40L-deficient cells and at an expanded level for wild-type CD4<sup>+</sup> T cells (Fig. 4j). To determine the stage at which the numbers of CD40L-deficient T cells fail to expand, adoptively transferred Cyt-c-specific TCR T cells were analysed for entry into the cell cycle. Wild-type TCR<sup>+</sup> T cells enter the cell cycle and substantial numbers of cells (35.5%) were seen in the S/M/

G2 phases of the cell cycles, with the remaining cells in G0/G1 phase (64.5%) (Fig. 4l), whereas most (89%) TCR<sup>+</sup> cells from CD40L-deficient mice were in G0/G1 phases, with few cells entering S/M/G2 phases (11%) (Fig. 4k). This suggests that the defect that leads to failure of the clonal expansion of T cells in CD40L-deficient mice is at an early stage of T-cell activation, preventing T cells from entering the cell cycle.

How then does the CD40-CD40L interaction regulate T-cell priming? This interaction is required for the induction of co-stimulatory activity in B cells<sup>8-10</sup>. We have recently confirmed that wild-type, but not CD40L-deficient T cells can activate co-stimulatory activity in B cells *in vitro*<sup>11</sup>. We believe, however, that this mechanism is unlikely to explain the failure of T-cell priming in CD40L-deficient mice. First, it has recently been shown using B-cell-deficient mice that B cells are not required for T-cell priming to KLH, one of the antigens used in our study<sup>12</sup>. Thus, CD40L activation of B-cell function would not be required for the T-cell clonal expansion we have studied, and therefore could not explain the deficiency in CD40L-deficient animals. Second, the same deficiency in T-cell priming is evident in immunization of myelin basic protein (MBP)-specific TCR transgenic mice with a short peptide of MBP which provokes experimental allergic encephalomyelitis (EAE) in control but not CD40L-deficient animals (I.S.G. and R.A.F., unpublished observations); T-cell response to peptide antigens requires dendritic cell but not B-cell APC function<sup>13</sup>. One possibility still to be investigated is that the CD40-CD40L interaction is necessary for the activation of co-stimulatory activity in APC such as dendritic cells, which are believed to initiate the immune response.

In summary, our results suggest that CD40L is required for *in vivo* clonal expansion of antigen-specific T cells, which in turn may explain the susceptibility of HIGM1 patients to *Cryptosporidial* and *Pneumocystis* infections. Likewise, CD40L-deficient mice have an impaired ability to resist *Leishmania* infection, which is probably contained by T-cell response leading to macrophage activation (L. Soong *et al.*, manuscript submitted). Agonists or antagonists of the CD40-CD40L interaction may eventually have therapeutic benefits: for example, an agent directed at the function of CD40L might be useful in the treatment of autoimmune diseases such as multiple sclerosis. □

Received 24 August; accepted 6 October 1995.

1. Kroccek, R. A. *et al.* *Immun. Rev.* **138**, 39-59 (1994).
2. Allen, R. C. *et al.* *Science* **259**, 990-993 (1993).
3. DiSanto, J. P., Bonnefoy, J. Y., Gauchat, J. F., Fischer, A. & de Saint Basile, G. *Nature* **361**, 541-543 (1993).
4. Callard, R. E., Armitage, R. J., Fanslow, W. C. & Spriggs, M. K. *Immun. Today* **14**, 559-564 (1993).
5. Notarangelo, L. D., Duse, M. & Ugazio, A. G. *Immunodef. Rev.* **3**, 101-121 (1992).
6. Xu, J. *et al.* *Immunity* **1**, 423-431 (1994).
7. Keamey, E. R., Pape, K. A., Loh, D. Y. & Jenkins, M. K. *Immunity* **1**, 327-339 (1994).
8. Kennedy, M. K. *et al.* *Eur. J. Immun.* **24**, 116-123 (1994).
9. Foy, M. *et al.* *Eur. J. Immun.* **25**, 596-603 (1995).
10. Ranheim, E. A. & Kipps, T. J. *J. exp. Med.* **177**, 925-935 (1993).
11. Wu, Y. *et al.* *Curr. Biol.* **5**, 1303-1311 (1995).
12. Epstein, M. M., Rosa, F. D., Jankovic, D., Sher, A. & Matzinger, P. *J. Exp. Med.* **182**, 915-922 (1995).
13. Constant, S. *et al.* *J. Immun.* **154**, 4915-4923 (1995).
14. Moudgil, K. D. & Sercarz, E. E. *J. exp. Med.* **178**, 2131-2138 (1993).
15. Kamogawa, Y., Minasi, L. A., Carding, S. R., Bottomly, K. & Flavell, R. A. *Cell* **75**, 985-995 (1993).
16. Strasser, A., Harris, A. W., Jacks, T. & Cory, S. *Cell* **79**, 329-339 (1994).

ACKNOWLEDGEMENTS. I.S.G. and J.X. are Associates and R.A.F. an Investigator of the Howard Hughes Medical Institute. I.S.G. and X.U. contributed equally to this project. We thank K. Bottomly for Cyt-c-specific-TCR-transgenic mice, T. Taylor for help with FACS analysis, and D. P. Grewal for technical assistance.

## CD40 ligand-transduced co-stimulation of T cells in the development of helper function

Dominic van Essen\*, Hltoshi Kikutani† & David Gray\*†

\* Department of Immunology, Royal Postgraduate Medical School, Hammersmith Hospital, Du Cane Road, London W12 0NN, UK

† Division of Immunology, Institute of Molecular and Cellular Biology, Osaka University, Osaka 565, Japan

MICE that lack either CD40<sup>1,2</sup> (expressed on B cells) or CD40 ligand<sup>3,4</sup> (expressed on activated T cells) are able neither to make IgG, IgA or IgE antibody responses, nor to generate germinal centres (the sites of formation of memory B cells). It has been assumed that these lesions were the result of an absence of signals to B cells through CD40. Here we show that the failure to signal T cells through CD40 ligand is an important contributory cause. Administration of soluble CD40 *in vivo* to CD40 knockout mice, restoring the missing signal through CD40 ligand, initiates germinal centre formation. Furthermore, T cells primed in the absence of CD40 (in CD40 knockout mice) are unable to help normal B cells to class switch or to form germinal centres (GC). These results indicate that co-stimulation of T cells through CD40 ligand causes their differentiation into cells that help B cells to make mature antibody responses and to generate memory populations.

To test the ability of T cells primed in the absence of CD40 to help B cells, we immunized CD40 knockout mice, and normal

littermates, with keyhole-limpet haemocyanin (KLH). Seven days later, purified T cells from these mice were injected into lightly irradiated CD40 knockout mice, with B cells from non-immune, CD40-expressing, *IgH<sup>a</sup>* congenic donors. These allotype-distinct 'indicator' B cells enable the donor/recipient source of serum antibodies to be determined. Adoptive hosts were immunized with dinitrophenylated KLH (DNP-KLH) (Fig. 1). Ten days after cell transfer, the mice that had received T cells primed in CD40<sup>+</sup> environment produced anti-DNP IgM, IgG1, IgG2a and IgG2b antibodies, and formed GC in their spleens. Mice that received T cells primed in CD40 knockout mice produced only IgM antibodies and exhibited no splenic GC (Fig. 1). The IgM and IgG antibodies were of donor *IgH<sup>a</sup>* allotype. The response to DNP-KLH is T-dependent as athymic, nude mice make no anti-DNP or anti-KLH antibodies (results not shown). Thus, T-dependent IgM production can proceed in the absence of CD40 signals to B cells<sup>1,2</sup> and without full-blown helper activity. As full helper activity does not develop over the 14 days in the presence of CD40<sup>+</sup> B cells in the adoptive hosts it seems that the CD40-CD40L interaction is crucial during the T-cell priming event and, possibly, cells other than B cells are required.

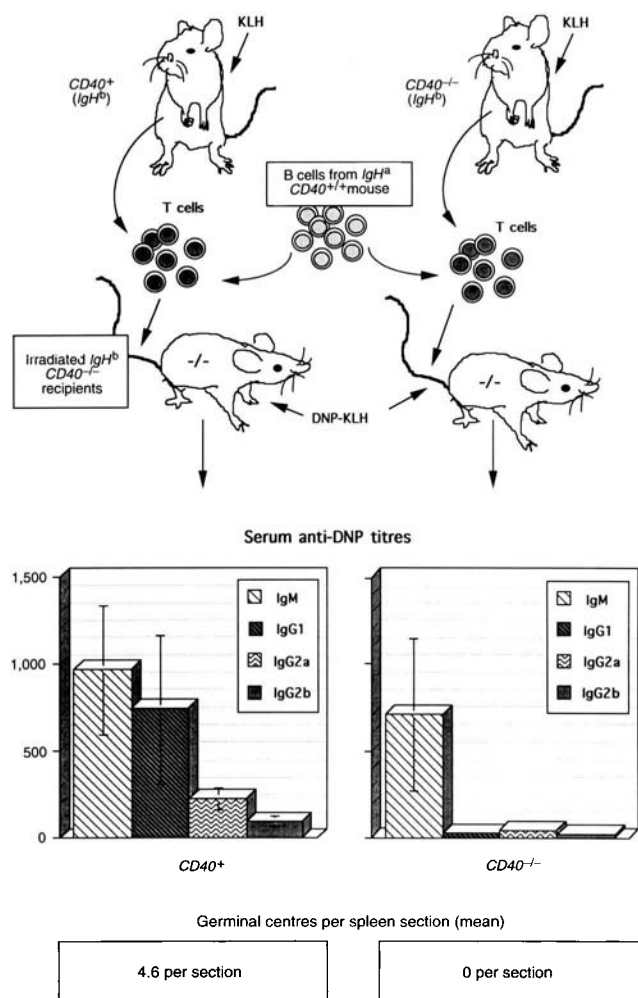
It is possible that T cells from CD40<sup>-/-</sup> mice are not tolerant of the CD40<sup>+</sup> donor B cells and might therefore reject them. We find no evidence for this as donor *IgM<sup>a</sup>* serum responses were made and *IgH<sup>a</sup>*-positive B cells could be detected in the spleens of recipient mice (results not shown). Likewise, the production of high titres of anti-DNP IgM antibodies indicates that the T cells have not been rendered unresponsive, as recently suggested after immunization of CD40-ligand knockout mice with allogeneic cells<sup>5</sup>.

Although the induction of T-dependent IgM antibody responses in CD40<sup>-/-</sup> mice<sup>1,2</sup> indicates effective T cell priming, we measured the antigen-specific T-cell precursor frequencies, after immunization of the knockout mice with KLH. This revealed variations between CD40<sup>-/-</sup> mice with a tendency

\* To whom correspondence should be addressed.

FIG. 1 Impaired helper function in T cells primed in a  $CD40^{-/-}$  environment. Irradiated  $CD40$  knockout mice were reconstituted with  $IgH^a$  B cells from normal, naive donors and T cells from normal or  $CD40$  knockout mice, previously primed with KLH. Recipient mice were immunized with DNP-KLH and 10 days later isotypes of serum anti-DNP antibodies were assayed and spleens examined for GC. IgM and IgG1 antibodies were of donor,  $IgH^a$  allotype. The experiment shown is representative of three experiments, involving four or five mice per group.

**METHODS.**  $CD40$  knockout<sup>1</sup> and C57BL/6C-IGA ( $IgH^a$  allotype congenic) mice were used, of 6–16 weeks of age.  $CD40^{-/-}$  and  $CD40^{+/-}$  or  $+/+$  mice were distinguished by FACS analysis of blood leucocytes stained with the anti- $CD40$  antibody FGK-45 (provided by J. Anderson and A. Rolink). Recipient mice were irradiated (5 Gy) 1 day before cell transfer. They received a mixture of  $2 \times 10^7$  B cells from C57BL/6C-IGA mice and  $5 \times 10^6$   $CD4^{+}$  T cells from either  $CD40^{-/-}$  or  $CD40^{+}$  mice. B cells were purified as described previously<sup>15</sup>.  $CD4^{+}$  T cells were enriched by depleting MHC Class II<sup>+</sup>,  $CD8^{+}$  and  $IgM^{+}$  cells from splenocytes, using biotinylated antibodies, M5-114 (American Type Culture Collection; ATCC), 53.6.7 (ATCC) and goat anti-mouse IgM (Southern Biotechnology Associates, Birmingham, AL) followed by MACS column separation (Miltenyi Biotech, Bergisch Gladbach, Germany). Mice were immunized with 100  $\mu$ g of alum-precipitated KLH or DNP-KLH intraperitoneally (i.p.). Immunoglobulin isotype levels were assayed by ELISA<sup>22</sup> and the host/donor origin of IgM and IgG1 was determined with  $IgH^a$ -specific Abs RS3.1 (ref. 23) and 20.9 (ref. 24). The values shown are the means  $\pm$  s.e.m. for four mice per group. Tissue sections were prepared and peanut-agglutinin (PNA)-FITC (fluorescein isothiocyanate) staining of GC was performed as described<sup>15</sup>. GC were enumerated by counting PNA<sup>+</sup> areas in sections through the whole spleen, at two different levels, separated by at least 50  $\mu$ m. No GC were observed in the spleens of recipients of T cells from  $CD40^{-/-}$  mice (mean per section,  $0 \pm 0$ ), compared with a mean of  $4.6 \pm 2.6$  per section in recipients of  $CD40^{+}$  T cells. All recipient mice, from two experiments (eight mice per group), were analysed and a mean figure of GC per spleen section was calculated.



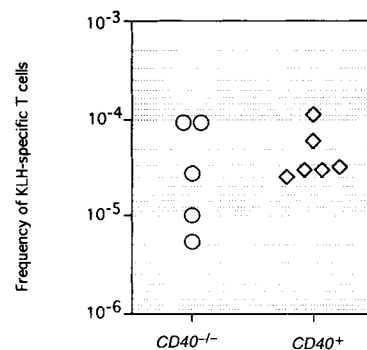
towards lower frequencies than in  $CD40^{+}$  mice (Fig. 2). Although general levels of T-cell priming in  $CD40$  knockout mice may be quantitatively impaired, this does not explain the qualitative defect in help for B-cells (Fig. 1): when the experiment shown in Fig. 1 was repeated such that the recipients of T cells primed in  $CD40^{-/-}$  mice received a 15-fold excess of antigen-specific T cells (measured retrospectively by LDA) over

recipients of T cells primed in  $CD40^{+}$  mice, the outcome was the same (results not shown).

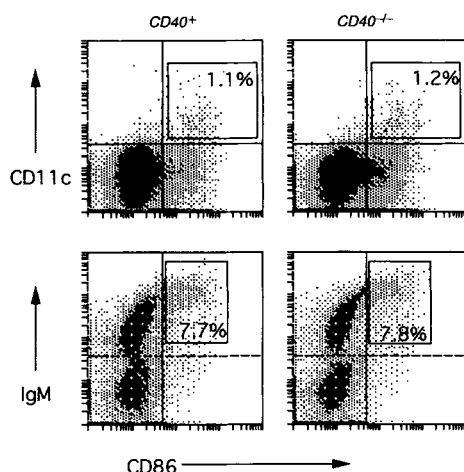
The failure to generate T-cell help might result from subnormal expression of the co-stimulatory molecules CD80 and CD86 on B cells and dendritic cells<sup>6-8</sup>. This does not seem to be the case as the number and proportion of CD86-expressing B cells and dendritic cells in the spleens of  $CD40^{+}$  and  $CD40^{-/-}$  mice

FIG. 2 Precursor frequencies of antigen-specific T cells in immunized  $CD40$  knockout and normal mice. Mice were immunized with KLH, and 10 days later the frequency of antigen-specific T cells was determined by limiting dilution analysis (LDA). The points represent the frequency of antigen-specific T cells in whole splenocyte populations; the proportion of T cells in the spleen of  $CD40^{+}$  and  $CD40^{-/-}$  mice was the same. The level of detection in this assay is  $\sim 1/200,000$ ; unimmunized  $CD40^{+}$  or  $CD40^{-/-}$  mice exhibited frequencies below this limit.

**METHODS.**  $CD40^{-/-}$  and  $CD40^{+}$  mice were immunized with 100  $\mu$ g alum-precipitated KLH i.p. The frequency of antigen-specific T cells in the spleen was determined by IL-2 LDA (ref. 25). T cells were depleted from splenocytes of unimmunized C57BL/6 mice, and irradiated (25 Gy);  $4 \times 10^4$  of these antigen-presenting cells were added to limiting numbers of responder splenocytes, with and without 50  $\mu$ g/ml KLH, in 96-well round-bottomed plates. At least 24 identical wells were set up at each concentration of responder cells. Two days later, plates were irradiated (25 Gy), and  $10^4$  cells of the IL-2-dependent CTLL cell line<sup>25</sup> (washed and re-cultured without IL-2 1 d before use) were added to each well. After a further 8 h, proliferation was assayed by uptake of [<sup>3</sup>H]-thymidine over 16 h. Positive wells exhibited a proliferation of greater than the mean plus 3 s.d. of control wells with no responder cells. The frequency of antigen-specific cells was calculated by regression analysis of the number of positive wells at each dilution of



responder cells. All points had a  $\chi^2$  value of  $<15$ . Before setting up the assay, the proportion of total and  $CD4^{+}$  T cells was assessed by FACS analysis of spleen cells stained with T24-FITC (anti-Thy1) or GK1.5-FITC (anti- $CD4$ ). We detected no differences in absolute number or proportion of total or  $CD4^{+}$  T cells between the  $CD40^{+}$  and  $CD40^{-/-}$  mice.



were similar (Fig. 3). The expression of CD86 on B cells may be induced by signalling via either surface immunoglobulin<sup>9-11</sup> or major histocompatibility complex (MHC) class II<sup>12,13</sup>. Although we favour a direct T-cell co-stimulatory role for CD40-ligand-transduced signals, the possibility remains that the CD40 signal to B cells or dendritic cells induces expression of, as yet, undefined co-stimulatory molecules that are important for the generation of helper function. It was previously reported that stimulation through CD40 ligand influences certain T-cell responses *in vitro*<sup>14</sup>.

To investigate directly CD40-ligand-transduced signals in T-cell maturation, we immunized CD40 knockout mice with DNP-KLH and injected them with a soluble CD40-Fc- $\gamma$ 1 fusion protein to crosslink the CD40 ligand. Ten days after immunization, mice treated with soluble CD40-Fc- $\gamma$ 1 had formed GC

(Fig. 4b), whereas mice injected with human IgG1 or with soluble CD80-Fc- $\gamma$ 1 fusion protein had not (Fig. 4a). A similar outcome was observed when CD40 knockout mice were crossed with transgenic mice that express soluble CD40 in the serum (P. Lane, personal communication). Unimmunized mice infused with CD40-Fc- $\gamma$ 1 do not significantly reconstitute the GC reaction. Interestingly, CD40 knockout mice treated with CD40-Fc- $\gamma$ 1 were still unable to make class-switched antibodies (Fig. 4e). This indicates that CD40 signalling to B cells, which is still absent in this experiment, is required for isotype switching but is not necessary for the initiation of the GC reaction. These data explain the discrepancy between experiments with CD40 and CD40-ligand knockout mice<sup>1-4</sup> (unable to form GC) and those with normal mice injected with CD40-Fc- $\gamma$ 1 to block the CD40-CD40-ligand interaction<sup>15</sup> (spares GC formation while

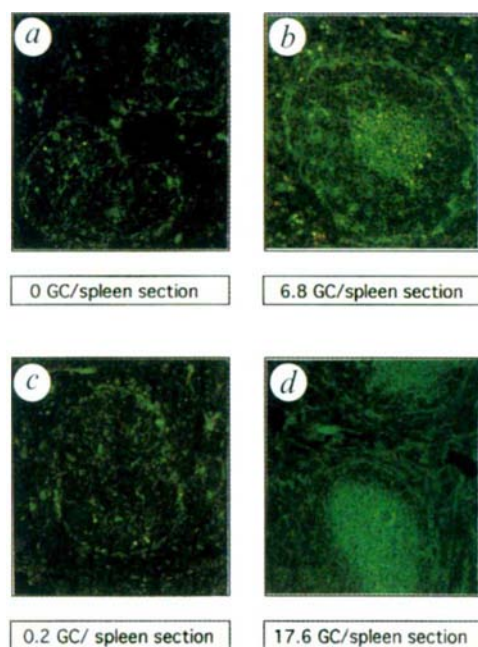


FIG. 4 Germinal centre formation in CD40 knockout mice treated with CD40-Fc- $\gamma$ 1 or CD80-Fc- $\gamma$ 1. Sections were cut from the spleens of CD40 knockout mice 10 days after immunization with DNP-KLH, and stained with PNA-FITC. a, GC do not normally form after immunization of CD40<sup>-/-</sup> mice, nor after infusion of soluble CD80-Fc- $\gamma$ 1 fusion protein (mean per section,  $0 \pm 0$ ). The green staining is of stromal elements that surround the splenic white pulp. Infusion of human IgG1 gave the same picture. The CD80-Fc- $\gamma$ 1 fusion protein was used as a control that would bind to T cells and mimic any nonspecific Fc-mediated effects on the accessory cells (for example, follicular dendritic cells) in B-cell

follicles. b, GC are regenerated when soluble CD40-Fc- $\gamma$ 1 fusion protein is injected (mean per section,  $6.8 \pm 2.1$ ). c, Infusion of CD40-Fc- $\gamma$ 1 fusion protein into unimmunized mice induces minimal GC formation (mean per section,  $0.2 \pm 0.0$ ); although the fusion protein carries foreign antigenic determinants it is administered in a relatively non-immunogenic soluble form. d, GC in CD40<sup>+</sup> mice after immunization with DNP-KLH (mean per section,  $17.6 \pm 4.0$ ). Counts of GC were carried out as described in the legend to Fig. 1; the data shown are the mean numbers per whole spleen section (six to ten mice per group). e, IgM and IgG antibody titres to DNP in the treated CD40<sup>-/-</sup> mice: all mice made anti-DNP IgM antibodies, but none of the CD40 knockout mice were able to produce anti-DNP IgG. There was no change in the absolute numbers or proportion of CD4<sup>+</sup> T cells in CD40-Fc- $\gamma$ 1 treated mice (results not shown).

METHODS. Purified human IgG1 myeloma protein (The Binding Site, Birmingham, UK), CD86-Fc- $\gamma$ 1 and CD40-Fc- $\gamma$ 1 were all injected at 150  $\mu$ g/d for 7 d after immunization (100  $\mu$ g alum-precipitated KLH, i.p.). Examination of tissue sections and measurement of serum antibody titres was performed as described in the legend to Fig. 1. Assessment of CD4<sup>+</sup> T cells was carried out as in the legend to Fig. 2.



inhibiting IgG). The recent observation that CD40L can also be expressed on B cells<sup>16</sup> means that we cannot be sure whether the effect of CD40-Fc- $\gamma$ 1 is due to CD40-ligand-transduced signals to T cells or to B cells. The fact that we observe a defect in T helper function in *CD40*<sup>-/-</sup> mice leads us to favour the notion of a signal to T cells. The transduction of signals to T cells by CD40 ligand could explain why some of the lesions in hyper-IgM syndrome patients indicate T-cell deficiency (for instance, *Pneumocystis* infections)<sup>17-20</sup>.

The GC formed after CD40-Fc- $\gamma$ 1 treatment of CD40 knock-out mice were underdeveloped: they were fewer and noticeably smaller than in *CD40*<sup>+</sup> littermates (Fig. 4b, d). It is possible that the CD40 signal to B cells (missing in these mice) is important for their progression within the GC. This may relate to direct effects of CD40 signals, inducing proliferation, or indirect ones, such as the upregulation of cytokine receptors<sup>21</sup>. We stress that the full maturation of germinal centres and the production of a full range of antibody isotypes seems to require both CD40 and CD40-ligand-mediated signals.

We do not know whether CD40 ligand co-stimulation plays a role in the clonal expansion of subsets of T cells or whether it regulates their differentiation and cytokine secretion. We do conclude that the interaction between B cells and T cells mediated through CD40 and CD40 ligand is bidirectional, enabling T cells to deliver effective help to B cells and allowing B cells to respond to it. □

Received 22 August; accepted 1 November 1995.

1. Kawabe, T. et al. *Immunity* **1**, 167-178 (1994).
2. Castigli, E. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 12135-12139 (1994).
3. Renshaw, B. R. et al. *J. exp. Med.* **180**, 1889-1900 (1994).
4. Xu, J. et al. *Immunity* **1**, 423-431 (1994).
5. Buhmann, J. E. et al. *Immunity* **2**, 645-653 (1995).
6. Ranheim, E. A. & Kipps, T. J. *J. exp. Med.* **177**, 925-935 (1993).
7. Kennedy, M. K. et al. *Eur. J. Immun.* **24**, 116-123 (1994).
8. Roy, M. et al. *Eur. J. Immun.* **25**, 596-603 (1995).
9. Freeman, G. J. et al. *J. Immun.* **143**, 2714-2722 (1989).
10. Lenschow, D. J. et al. *J. Immun.* **153**, 1990-1997 (1994).
11. Ho, W. Y., Cooke, M. P., Goodnow, C. C. & Davis, M. M. *J. exp. Med.* **179**, 1539-1549 (1994).
12. Koulouva, L., Clark, E. A., Shu, G. & Dupont, B. *J. exp. Med.* **173**, 759-762 (1991).
13. Nabavi, N. et al. *Nature* **360**, 266-268 (1992).
14. Cayabyab, M., Phillips, J. H. & Lanier, L. J. *Immun.* **152**, 1523-1531 (1994).
15. Gray, D., Dullforce, P. & Jainandunsing, S. *J. exp. Med.* **180**, 141-155 (1994).
16. Grammer, A. C. et al. *J. Immun.* **154**, 4996-5010 (1995).
17. Aruffo, A. et al. *Cell* **72**, 291-300 (1993).
18. DiSanto, J. P., Bonnefoy, J. Y., Gauchat, J. F., Fischer, A. & de Saint Basile, G. *Nature* **361**, 541-543 (1993).
19. Fuleihan, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 2170-2173 (1993).
20. Korthäuer, U. et al. *Nature* **361**, 539-541 (1993).
21. Gray, D., Siepmann, K. & Wöhlleben, G. *Semin. Immun.* **6**, 303-310 (1994).
22. Gray, D., MacLennan, I. C. M. & Lane, P. J. L. *Eur. J. Immun.* **16**, 641-648 (1986).
23. Schüppel, R., Wilke, J. & Weiler, E. *Eur. J. Immun.* **17**, 739-741 (1987).
24. Oi, V. T. & Herzenberg, L. A. *Molec. Immun.* **16**, 1005-1017 (1979).
25. Ronchese, F. & Hausmann, B. *J. exp. Med.* **177**, 679-690 (1993).
26. Kosco, M., Pflugfelder, E. & Gray, D. (1992). *J. Immun.* **148**, 2331-2339 (1992).
27. Metlay, J. P. et al. *J. exp. Med.* **171**, 1753-1771 (1990).
28. Hathcock, K. S. et al. *Science* **262**, 905-907 (1993).

ACKNOWLEDGEMENTS. We thank P. Lane (Basel Institute for Immunology) for his gifts of fusion proteins and for permission to quote his unpublished data. We thank P. Lane, B. Stockinger, T. Zai and R. Lechler for discussions and critical review of the manuscript. This work was supported by the Wellcome Trust and the Medical Research Council (UK).

## Functional impact of syntaxin on gating of N-type and Q-type calcium channels

I. Bezprozvanny\*, R. H. Scheller† & R. W. Tsien‡

Department of Molecular and Cellular Physiology,  
†Howard Hughes Medical Institute, Beckman Center,  
Stanford University Medical Center, Stanford, California 94305, USA

RAPID and reliable synaptic transmission depends upon the close proximity of voltage-gated calcium channels and neurotransmitter-containing vesicles in the presynaptic terminal<sup>1</sup>. Although it is clear that a local  $\text{Ca}^{2+}$  rise conveys the crucial signal from  $\text{Ca}^{2+}$  channels to the exocytotic mechanism<sup>2</sup>, little is known about whether communication ever proceeds in the opposite direction, from the release machinery to  $\text{Ca}^{2+}$  channels<sup>3,4</sup>. To look for such signalling, we examined the interaction of various types of voltage-gated  $\text{Ca}^{2+}$  channels with syntaxin, a presynaptic membrane protein of relative molecular mass 35,000 (refs 5-7) which may play a key part in synaptic vesicle docking and fusion<sup>8</sup> and which interacts strongly with N-type  $\text{Ca}^{2+}$  channels<sup>5,6,9-12</sup>. Here we report that co-expression of syntaxin 1A with N-type channels in *Xenopus* oocytes sharply decreases the availability of these channels. This is due to the stabilization of channel inactivation rather than to a simple block or lack of channel expression, because it is overcome by strong hyperpolarization. Deletion of syntaxin's carboxy-terminal transmembrane domain abolishes its functional effect on  $\text{Ca}^{2+}$  channels. Syntaxin produced a similar effect on Q-type  $\text{Ca}^{2+}$  channels encoded by  $\alpha_{1A}$  (refs 13-15) but not on L-type  $\text{Ca}^{2+}$  channels. Thus, the syntaxin effect is specific for  $\text{Ca}^{2+}$  channel types that participate in fast transmitter release in the mammalian central nervous system (ref. 16 and therein). We hypothesize that, in addition to acting as a vesicle-docking site, syntaxin may influence presynaptic  $\text{Ca}^{2+}$  channels, opposing  $\text{Ca}^{2+}$  entry where it is not

advantageous, but allowing it at release sites where synaptic vesicles have become docked and/or ready for fusion.

Co-expression of syntaxin 1A with N-type channels left the peak current-voltage relationship unchanged (data not shown), but significantly altered the balance between channel availability and inactivation (Fig. 1). A 'descending staircase' voltage protocol was used to compare the behaviour of N-type channels ( $\alpha_{1B}\beta_3\alpha_2$ ), expressed by themselves (control in Fig. 1a), co-expressed with full-length syntaxin 1A (s1A in Fig. 1b), or co-expressed with the N-terminal two-thirds of syntaxin 1A (s1A-N in Fig. 1c). s1A-N complementary-RNA was injected in the same concentration as s1A cRNA (~20 ng per oocyte) and provided a negative control as it does not interact with N-type  $\text{Ca}^{2+}$  channels<sup>12</sup>. In contrast to either control, coexpression of full-length s1A consistently decreased the availability of N-type channels at -60 and -80 mV holding potentials. Peak currents evoked from -80 mV ( $I_{80}$ ) were normalized to those evoked from -120 mV ( $I_{120}$ ). The  $I_{80}/I_{120}$  ratio averaged  $0.47 \pm 0.02$  ( $n = 38$ ) for control oocytes,  $0.11 \pm 0.01$  ( $n = 33$ ) with co-expression of s1A, and  $0.52 \pm 0.01$  ( $n = 12$ ) with co-expression of s1A-N. The effect of s1A cannot be attributed to the mere presence of foreign protein in the surface membrane, as high-level co-expression of platelet-activating-factor receptor<sup>17</sup> did not affect N-type channel gating ( $I_{80}/I_{120} = 0.46 \pm 0.06$  ( $n = 4$ )). The four-fold smaller value of  $I_{80}/I_{120}$  in the presence of syntaxin 1A suggests that s1A significantly favours the inactivated state of the channel. Co-expression of s1A also slowed the recovery from inactivation upon hyperpolarization (Fig. 1, bottom). On hyperpolarization to 120 mV, exponential time constants of recovery ( $\tau_{120}$ ) were  $11.7 \pm 0.03$  s ( $n = 38$ ) for control oocytes,  $26.8 \pm 1.7$  s ( $n = 33$ ) with s1A, and  $11.2 \pm 0.4$  s ( $n = 12$ ) with s1A-N—a more than 2-fold slowing with s1A relative to the controls. Increases in  $\tau_{120}$  and decreases in  $I_{80}/I_{120}$  were significantly correlated ( $r = 0.75$ ), which is consistent with a common mechanism for both changes.

Oocytes already expressing N-type  $\text{Ca}^{2+}$  channels were re-injected with cRNA encoding s1A to see whether alterations in channel properties grew with time, as would be expected for progressively increasing s1A expression. Injection of s1A cRNA was followed by a gradually developing decrease in  $I_{80}/I_{120}$  (Fig. 2a) and an increase in  $\tau_{120}$  (Fig. 2b), whereas injection of water

\* Permanent address: Institute of Cytology, Russian Academy of Sciences, St Petersburg, 194064 Russia.

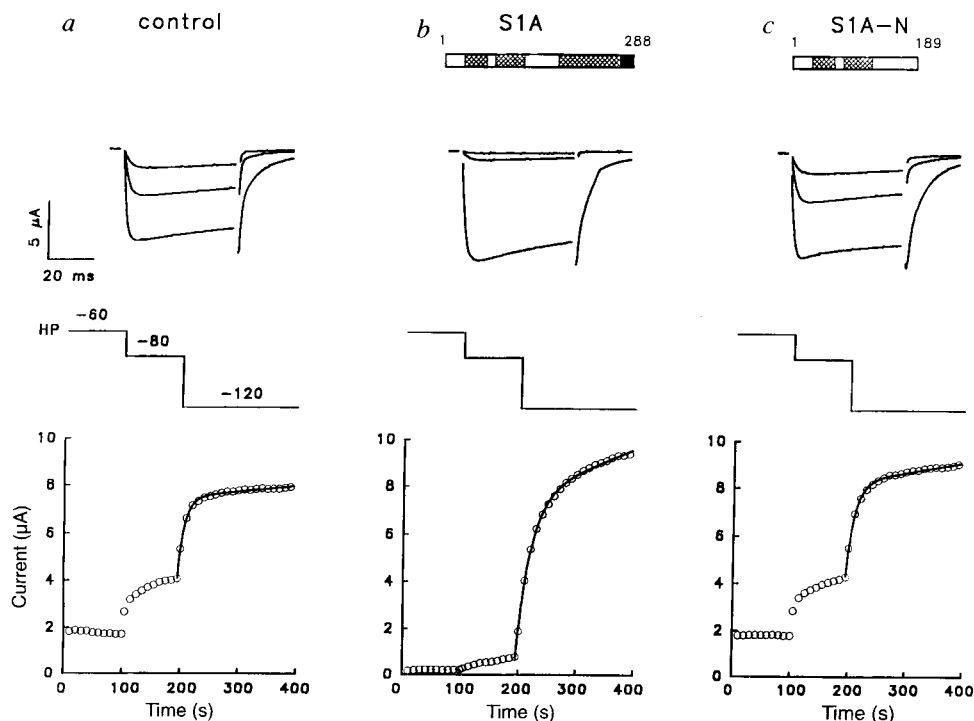
‡ To whom correspondence should be addressed.

FIG. 1 Functional properties of N-type channels in *Xenopus* oocytes, expressed a, alone, or b, co-expressed with full-length s1A (amino acids 1–288), or c, with the N-terminal portion of s1A (s1A-N, amino acids 1–189). Domain structure of syntaxin (top) includes putative helical domains that have a propensity to interact with other proteins<sup>19,20</sup> (shaded) and a putative transmembrane domain<sup>7</sup> (filled). Inward  $\text{Ba}^{2+}$  currents were recorded with a 'descending staircase' stimulation protocol. 50-ms test pulses to 0 mV (the peak of the  $I$ - $V$  relationship)<sup>26,27</sup> were applied at 0.1 Hz as the holding potential (HP) was switched from  $-60$  mV (10 pulses) to  $-80$  mV (10 pulses), and then to  $-120$  mV (20 pulses). Current traces are the last sweep at each HP. Bottom, peak current amplitudes plotted against time. The smooth curve fits the increase in peak current after transition to HP =  $-120$  mV and obeys the equation  $I(t) = I_{80} + C(1 + at)[1 - \exp(-t/\tau_{120})]$ . The linear term corrects for very slow removal of inactivation at strongly negative potentials, not studied in detail here; without this correction, estimates of syntaxin-induced changes in  $\tau_{120}$  would be even larger.

**METHODS.** The coding sequence of rat s1A (ref. 5) was isolated by polymerase chain reaction (PCR) and introduced into the pGEMHE *Xenopus* oocyte expression vector<sup>18</sup>. The N-terminal construct (s1A-N) was generated in the same way, but with a stop codon replacing the triplet encoding Gln 190. Capped cRNA was synthesized *in vitro* as described<sup>27</sup> and tested by translation in a rabbit reticulocyte lysate system before injection into oocytes. The efficiency of translation was monitored by incorporation of  $^{35}\text{S}$ -Met using a FUJIBAS-2000 phosphorimager.  $\text{Ca}^{2+}$  channel cRNA was synthesized with cDNA clones encoding human brain  $\alpha_{1B}$  (ref. 26), rabbit skeletal muscle  $\alpha_2/\delta$  (ref. 28), and rabbit brain  $\beta_3$

(control) or s1A-N failed to affect channel availability (as in Fig. 1). The effect of s1A became maximal  $\sim 2$  days after cRNA injection, but some effect was already apparent at 14 hr, consistent with rapid protein expression obtained with the pGEMHE vector in *Xenopus* oocytes<sup>18</sup>. The  $I_{80}/I_{120}$  ratio and  $\tau_{120}$  also showed a clear dependence on the amount of s1A cRNA introduced by reinjection (Fig. 2c, d). The concentration-dependent decrease in  $I_{80}/I_{120}$  ratio followed the Hill equation (Fig. 2c), whereas  $\tau_{120}$  increased linearly with the amount of s1A cRNA (Fig. 2d), as expected in a modulated receptor scheme in which syntaxin binds preferentially to the channel's inactivated state.

Although N-type  $\text{Ca}^{2+}$  channels encoded by  $\alpha_{1B}$  are the only channels known so far to interact with syntaxin<sup>5,6,9–12</sup>, they are just one of the  $\text{Ca}^{2+}$  channel types found at presynaptic terminals. Channels encoded by  $\alpha_{1A}$  subunits (Q- and possibly P-type)<sup>14,15</sup> are major mediators of  $\text{Ca}^{2+}$  entry and triggered transmitter release (see ref. 16 for references). L-type channels are present at nerve terminals, but play no detectable role in evoked synaptic transmission. Accordingly, we looked for possible responsiveness of Q- and L-type channels to syntaxin co-expression (Fig. 3b, c). Oocytes expressing  $\alpha_{1A}\beta_3\alpha_2$  were reinjected with water (control), or with cRNAs encoding s1A or s1A-N; inactivation properties of the resulting Q-type currents were evaluated 4–5 days later, using a protocol appropriate to the voltage-dependence of Q-type channels (Fig. 3b). The ratio of peak current amplitudes  $I_{60}/I_{120}$  was  $0.53 \pm 0.03$  ( $n=7$ ) in control oocytes,  $0.17 \pm 0.05$  ( $n=10$ ) with s1A, and  $0.57 \pm 0.07$  ( $n=5$ ) with s1A-N. Thus, Q-type channels are very similar to N-type channels in their responsiveness to syntaxin co-expression. L-



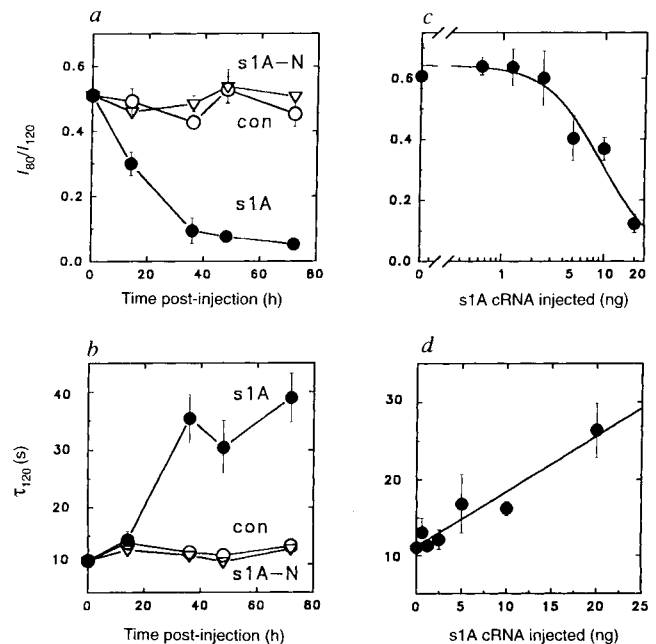
(ref. 29) as templates.  $\text{Ca}^{2+}$  channel subunit cRNAs were premixed in roughly equimolar ratios and injected into freshly isolated *Xenopus laevis* oocytes. 4–5 days later, oocytes were divided into several groups and reinjected with s1A cRNA, s1A-N cRNA or an equal volume of water (control). Expressed  $\text{Ca}^{2+}$  channel currents were evaluated 2–4 days after the second cRNA injection unless indicated otherwise. Two-microelectrode voltage clamp recordings were carried out with 5 mM external  $\text{Ba}^{2+}$  as charge carrier<sup>27</sup>. Leak and capacitance currents were subtracted on line with a P/6 protocol.

type channels ( $\alpha_{1C}\beta_3\alpha_2$ ) failed to respond in any detectable way to co-injection of s1A cRNA (Fig. 3c), in agreement with biochemical data<sup>11</sup>. The value of  $I_{80}/I_{100}$  with s1A ( $0.59 \pm 0.02$ ,  $n=20$ ) was no different from that in the control ( $0.63 \pm 0.02$ ,  $n=11$ ) or with s1A-N ( $0.59 \pm 0.03$ ,  $n=12$ ). Strikingly, the pattern of responsiveness of the channel types to syntaxin matches their participation in evoked neurotransmission at synapses of the mammalian central nervous system.

We examined the action of syntaxin over a wide range of holding potentials to see whether the inhibitory effects on N- and Q-type channels were accentuated as these channels became increasingly inactivated. If this were the case, the voltage-dependence of channel availability would be expected to shift towards more negative potentials. Figure 3d compares the voltage-dependence of N-type channels in oocytes reinjected with water (control), or with cRNAs for s1A or s1A-N. The midpoint voltage ( $V_{1/2}$ ) was  $-83.9 \pm 3.3$  mV ( $n=3$ ) in control,  $-104.6 \pm 2.8$  mV ( $n=6$ ) with s1A, and  $-86.5 \pm 2.5$  mV ( $n=4$ ) with s1A-N. Thus, full-length syntaxin 1A caused a  $\sim 20$  mV hyperpolarizing shift in the voltage dependence of N-type channel availability. Likewise, for Q-type channels, the estimates of  $V_{1/2}$  were  $-63.0 \pm 1.5$  mV ( $n=4$ ) in control,  $-83.2 \pm 3.1$  mV ( $n=3$ ) with s1A, and  $-62.1 \pm 1.4$  mV ( $n=4$ ) with s1A-N (Fig. 3e). Similar results were obtained for both N- and Q-type channels with co-expression of syntaxin 1B.

To narrow down the structural determinants of syntaxin's interaction with  $\text{Ca}^{2+}$  channels, we constructed a series of s1A mutants and evaluated their impact on N-type channels (data not shown). A C-terminal segment (amino acids 168–288), previ-

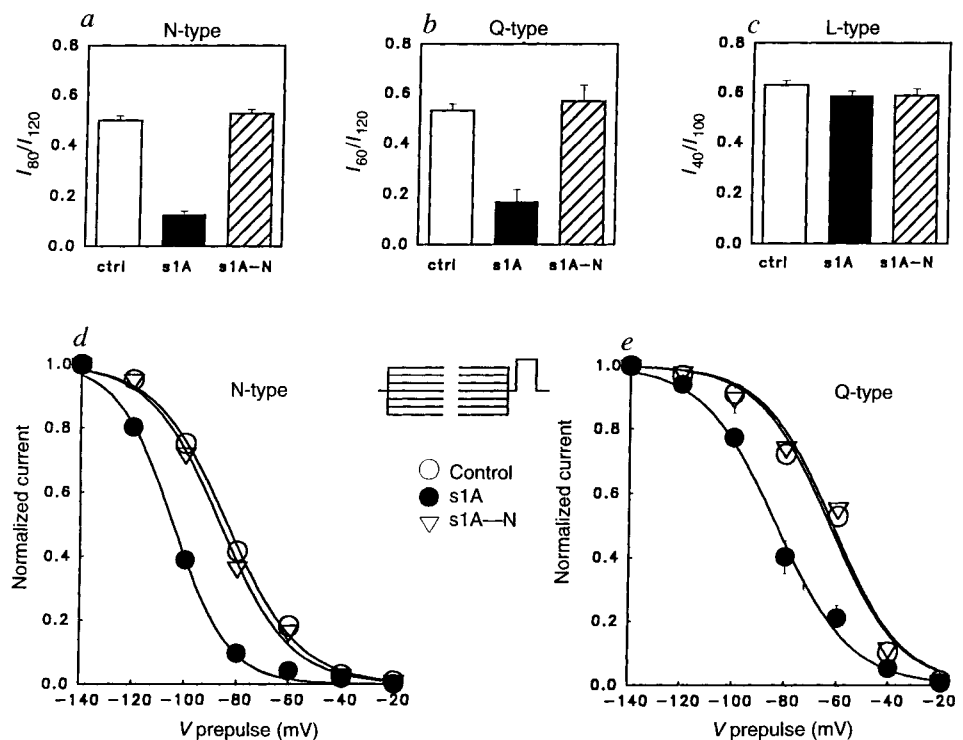
FIG. 2 Progressive changes in the functional properties of N-type  $\text{Ca}^{2+}$  channels with increasing duration of syntaxin expression and with increasing amount of syntaxin cRNA injected. *a*, *b*, The  $I_{80}/I_{120}$  ratio (*a*) and  $\tau_{120}$  (*b*) were determined using the descending staircase protocol at different times after injection of cRNA for syntaxin (filled circles), for the N-terminal construct s1A-N (triangles), or an equal volume of water (open circles). Each point represents mean value  $\pm$  s.e.m. ( $n=3$ ). The zero time point corresponds to data obtained just before syntaxin cRNA injection. *c*, *d*, The  $I_{80}/I_{120}$  ratio (*c*) and  $\tau_{120}$  (*d*) were determined using the descending staircase protocol in oocytes injected with similar volumes of s1A cRNA-containing solution, either undiluted ( $\sim 20$  ng per oocyte) or diluted 2-, 4-, 8-, 16- or 32-fold. Zero concentration corresponds to data obtained with a control batch of oocytes injected with water. Experiments were done four days after cRNA injection. Each point represents mean  $\pm$  s.e.m. ( $n=4-6$ ). The  $I_{80}/I_{120}$  data were fitted by least-squares to the smooth curve, which conforms to the Hill equation,  $I_{80}/I_{120} = 1 / (1 + (K^n / (K^n + c^n)))$ , where the parameters were  $n=1.6$ ,  $K=9.5$  ng per oocyte. One possible interpretation of these data is that more than one syntaxin molecule may bind to each N-type channel and contribute to reducing its availability. But more information is required about the relationship between cRNA concentration and the degree of protein translation in oocytes, and any effects of syntaxin expression on the number of N-type channels present in the membrane.



ously shown to interact with N-type channel sequence in biochemical experiments<sup>12</sup>, had nearly the same inhibitory effect on channel availability as full-length syntaxin itself. This segment incorporates a putative helical domain, H3 (amino acids 191–265), important for interactions with several key presynaptic proteins<sup>19–22</sup>. However, H3 did not appear to be the unique site

of syntaxin-channel interaction, as an s1A construct lacking this domain showed enhanced ability to modify channel properties ( $I_{80}/I_{120} = 0.042 \pm 0.009$  ( $n=21$ )). In contrast, deletion of syntaxin's C-terminal transmembrane domain (TMR, amino acids 266–288) completely eliminated the modulatory action. Although the inclusion of a TMR appeared to be essential, there

FIG. 3 Specificity of syntaxin effect on various types of  $\text{Ca}^{2+}$  channels. *a*–*c*, Comparison of the behaviour of N-type channels (*a*), Q-type channels (*b*) and L-type channels (*c*). Properties of N-type channels were studied using the descending staircase protocol (Fig. 1). Inactivation of Q- and L-type channels takes place over a less negative range of potentials than for N-type channels<sup>27</sup>; accordingly, the descending staircase protocol was changed from 60/80/120 to 40/60/120 for Q-type channels and to 40/100 for L-type channels. The peak currents for most negative holding potential ranged between 5–15  $\mu\text{A}$  for N-type, 5–25  $\mu\text{A}$  for Q-type, and 0.5–2.5  $\mu\text{A}$  for L-type. *d*, *e*, Inactivation curves of N-type and Q-type  $\text{Ca}^{2+}$  channels are shifted in the presence of syntaxin. Voltage-dependent inactivation curves were assessed with a prepulse protocol (inset). The oocyte membrane was held for 30 s at various potentials ( $V_{\text{prepulse}}$ ) between  $-20$  mV and  $-140$  mV, returned to  $-80$  mV for 20 ms, and then depolarized with a 50-ms test pulse to 0 mV to evoke inward  $\text{Ba}^{2+}$  current. For each oocyte, the peak current amplitudes associated with various levels of  $V_{\text{prepulse}}$  were normalized by the peak current with  $V_{\text{prepulse}} = -140$  mV to facilitate pooling of data from different cells within an experimental group. The normalized current magnitudes (mean  $\pm$  s.e.m.,  $n=3-6$ ) are plotted against  $V_{\text{prepulse}}$ . Smooth curves represent best fits to a Boltzmann relationship,  $I(V_{\text{prepulse}})/I(-140) = 1 / (1 + \exp((V_{\text{prepulse}} - V_{1/2})/k))$ . Values of the midpoint voltage  $V_{1/2}$  are given in the text. **METHODS.** cRNA for Q-type channels was made with an  $\alpha_{1A}$  cDNA template derived from rabbit brain<sup>30</sup>. cRNA for L-type channels was made with an  $\alpha_{1C}$  cDNA clone derived from rabbit heart<sup>28</sup>. Compared with N- and Q-type channels, L-type  $\text{Ca}^{2+}$ -channel currents are smaller



and express more slowly. Accordingly, oocytes expressing L-type currents were given more time (5–8 days) after injection of channel subunit cRNAs to allow further channel expression before reinjection with s1A and s1A-N cRNA. Recordings from L-type channels were made using 40 mM external  $\text{Ba}^{2+}$  solution and a test potential of  $+20$  mV. In some experiments, FPL 64176 ( $2 \mu\text{M}$ ) was present to enhance the L-type currents. Results obtained in the presence and absence of FPL 64176 were not significantly different and were averaged together.

was considerable flexibility in its amino-acid sequence. Replacement of syntaxin 1A TMR with TMRs from several other members of the syntaxin protein family<sup>7</sup>, or even a random rearrangement of the s1A TMR sequence, had only minor effects on the ability of the resulting constructs to modify channel gating. Additional experiments, with procedures to quantify the level of surface expression of syntaxin mutants in oocytes may further clarify the structural basis of the syntaxin-channel interactions.

We have shown that syntaxin has a functional impact on the gating of  $\text{Ca}^{2+}$  channels. It reduces the availability of N-type  $\text{Ca}^{2+}$  channels and slows their recovery from inactivation. These effects are likely to arise from direct syntaxin-N-type channel interaction<sup>5,6,9,12</sup>, but involvement of intermediary components in the oocyte expression system cannot be ruled out. Syntaxin's ability to modify gating behaviour is channel-specific, extending to Q-type but not L-type channels. For both N- and Q-type channels, we found that the voltage-dependence of inactivation was shifted by  $-20$  mV, greatly reducing the channel availability at a neuronal resting potential near  $-75$  mV. We therefore propose a new function for the syntaxin- $\text{Ca}^{2+}$  channel interaction, beyond that of keeping  $\text{Ca}^{2+}$  channels and presynaptic vesicles in close proximity<sup>5,6,9,12</sup>. Assuming that our findings are applicable to nerve terminals, where syntaxin is very abundant<sup>23</sup> and interacts with  $>70\%$  of N-type channels<sup>11</sup>, syntaxin would be expected strongly to inhibit presynaptic  $\text{Ca}^{2+}$  entry (see ref. 24). The inhibition would produce the desirable effect of reducing  $\text{Ca}^{2+}$  influx in the absence of docked synaptic vesicles. We speculate further that this inhibition might be relieved following the docking of a synaptic vesicle to syntaxin. Possible mediators of such disinhibition include exocytotic components that interact directly with syntaxin<sup>19,22</sup> (such as VAMP/synaptobrevin, synaptotagmin, SNAP-25, *n*-sec1 and  $\alpha$ -SNAP), as well as cysteine string proteins, which are closely associated with synaptic vesicles<sup>3,4</sup> and have a stimulatory effect on  $\text{Ca}^{2+}$  channel activity<sup>25</sup>. The hypothesized mechanism might signal the readiness of the exocytotic machinery for activation, allowing rapid  $\text{Ca}^{2+}$  entry precisely where it would be most appropriate.

**Note added in proof:** Recent biochemical experiments indicate that syntaxin 1A binds to  $\alpha_{1A}$   $\text{Ca}^{2+}$  channel proteins (Z. H. Sheng, J. Rettig and W. A. Catterall, personal communication; M. Seagar and M. Takahashi, personal communication), consistent with our finding of a functional interaction with Q-type channels. □

Received 14 June; accepted 17 October 1995.

1. Llinas, R., Steinberg, I. Z. & Walton, K. *Biophys. J.* **33**, 323–351 (1981).
2. Katz, B. *The Release of Neural Transmitter Substances* (Thomas, Springfield, Illinois, 1969).
3. Mastrogiacomio, A. et al. *Science* **263**, 981–982 (1994).
4. Zinsmaier, K. E., Eberle, K. K., Buchner, E., Walter, N. & Benzer, S. *Science* **263**, 977–980 (1994).
5. Bennett, M. K., Calakos, N. & Scheller, R. H. *Science* **257**, 255–259 (1992).
6. Yoshida, A. et al. *J. Biol. Chem.* **267**, 24925–24928 (1992).
7. Bennett, M. K. et al. *Cell* **74**, 863–873 (1993).
8. Sollner, T., Bennett, M. K., Whiteheart, S. W., Scheller, R. H. & Rothman, J. E. *Cell* **75**, 409–418 (1993).
9. Horikawa, H. P. M. et al. *FEBS Lett.* **330**, 236–240 (1993).
10. Leveque, C. et al. *J. Biol. Chem.* **269**, 6306–6312 (1994).
11. el Far, O. et al. *FEBS Lett.* **361**, 101–105 (1995).
12. Sheng, Z. H., Rettig, J., Takahashi, M. & Catterall, W. A. *Neuron* **13**, 1303–1313 (1994).
13. Randall, A. & Tsien, R. W. *J. Neurosci.* **15**, 2995–3012 (1995).
14. Sather, W. A. et al. *Neuron* **11**, 291–303 (1993).
15. Stea, A. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 10576–10580 (1994).
16. Wheeler, D. B., Tsien, R. W. & Randall, A. *Science* **266**, 830–831 (1994).
17. Bito, H., Honda, Z., Nakamura, M. & Shimizu, T. *Eur. J. Biochem.* **221**, 211–218 (1994).
18. Liman, E. R., Tytgat, J. & Hess, P. *Neuron* **9**, 861–871 (1992).
19. Chapman, E. R., An, S., Barton, N. & Jahn, R. *J. Biol. Chem.* **269**, 27427–27432 (1994).
20. Kee, Y., Lin, R. C., Hsu, S. C. & Scheller, R. H. *Neuron* **14**, 991–998 (1995).
21. McMahon, H. T. & Sudhof, T. C. *J. Biol. Chem.* **270**, 2213–2217 (1995).
22. Hayashi, T., Yamasaki, S., Nauenburg, S., Binz, T. & Niemann, H. *EMBO J.* **14**, 2317–2325 (1995).
23. Walch-Solimena, C. et al. *J. Cell. Biol.* **128**, 637–645 (1995).
24. Smirnova, T., Fossier, P., Stinnakre, J., Mallet, J. & Baux, G. *Neuroscience* **68**, 125–133 (1995).
25. Gundersen, C. B. & Umbach, J. A. *Neuron* **9**, 527–537 (1992).
26. Ellinor, P. T., Zhang, J.-F., Horne, W. A. & Tsien, R. W. *Nature* **372**, 272–275 (1994).
27. Bezprozvanny, I. & Tsien, R. W. *Molec. Pharmacol.* **48**, 540–549 (1995).
28. Mikami, A. et al. *Nature* **340**, 230–233 (1989).

29. Hullin, R. et al. *EMBO J.* **11**, 885–890 (1992).
30. Mori, Y. et al. *Nature* **350**, 398–402 (1991).

**ACKNOWLEDGEMENTS.** We thank H. Bito, X.-H. Chen and S.-C. Hsu for help with molecular biology techniques; R. Agin for technical assistance; S. Bezprozvannaya for support and encouragement; T. Tanabe, Y. Mori, F. Hofmann, V. Flockert, J.-F. Zhang and P. T. Ellinor for  $\text{Ca}^{2+}$ -channel subunit clones; Y. Kee for syntaxin point mutants; E. Liman for the pGEMHE vector; A. J. G. Baxter (Fisons PLC) for FPL 64176; H. Bito for help in experiments with PAF receptor cDNA; W. A. Catterall, M. Seagar and M. Takahashi for access to unpublished work; and H. Bito, X.-H. Chen, P. T. Ellinor, S.-C. Hsu, G. Liu, B. A. Premack, T. L. Schwarz, D. B. Wheeler and J.-F. Zhang for discussions and for comments on the manuscript. This work was supported by the NIMH-Silvio Conte Center at Stanford and by grants from the NIH, the G. Harold and Leila Y. Mathers Foundation (R.W.T. and R.H.S.), and a Cardiac and Cellular Electrophysiology training grant from the NIH (I.B.). R.H.S. is an HHMI Investigator.

## A phosphate transporter from the mycorrhizal fungus *Glomus versiforme*

Maria J. Harrison & Marianne L. van Buuren

The Samuel Roberts Noble Foundation, Plant Biology Division, 2510 Sam Noble Parkway, Ardmore, Oklahoma 73401, USA

VESICULAR-arbuscular (VA) mycorrhizal fungi form symbiotic associations with the roots of most terrestrial plants, including many agriculturally important crop species<sup>1–3</sup>. The fungi colonize the cortex of the root to obtain carbon from their plant host, while assisting the plant with the uptake of phosphate and other mineral nutrients from the soil<sup>2–5</sup>. This association is beneficial to the plant, because phosphate is essential for plant growth and development, especially during growth under nutrient-limiting conditions<sup>2,6–8</sup>. Molecular genetic studies of these fungi and their interaction with plants have been limited owing to the obligate symbiotic nature of the VA fungi, so the molecular mechanisms underlying fungal-mediated uptake and translocation of phosphate from the soil to the plant remain unknown. Here we begin to investigate this process by identifying a complementary DNA that encodes a transmembrane phosphate transporter (GvPT) from *Glomus versiforme*, a VA mycorrhizal fungus. The function of the protein encoded by GvPT was confirmed by complementation of a yeast phosphate transport mutant. Expression of GvPT was localized to the external hyphae of *G. versiforme* during mycorrhizal associations, these being the initial site of phosphate uptake from the soil.

A cDNA clone (GvPT) showing hybridization to the *PHO84* sequence from *Saccharomyces cerevisiae*<sup>9</sup> was isolated from a cDNA library prepared from roots of *Medicago truncatula* colonized with the mycorrhizal fungus, *G. versiforme*. GvPT is 1,932 base pairs (bp) in length, and the open reading frame (ORF) is flanked by 112 bp of untranslated sequence at the 5' end and 257 bp of untranslated sequence including the poly(A)<sup>+</sup> tail at the 3' end (GenBank accession No. U38650). The encoded protein of 521 amino acids is predicted to be an integral membrane protein containing 12 membrane-spanning domains which is a secondary structure shared by many membrane transporters from both eukaryotes and prokaryotes<sup>10,12</sup> (Fig. 1). Genes encoding phosphate transporters have been cloned from yeast (*PHO84*)<sup>9</sup> and *Neurospora crassa* (*PHO-4* and *PHO-5*)<sup>13,14</sup>. GvPT shares 47.9% sequence identity with *PHO84* and 45% identity with *PHO-5* (Fig. 1). *PHO-4* is considerably different from the other transporters and shares little sequence identity with either *PHO84*, *PHO-5* or GvPT. Note that the domains most conserved between these three proteins include sequences containing protein kinase C and casein kinase II phosphorylation sites (Fig. 1). Post-translational regulation by phosphorylation has been previously reported for glucose transporters<sup>15</sup>, but it is not yet known whether the phosphate transporters are regulated in this manner.

GvPT was identified from a cDNA library containing representatives of two genomes (plant and fungal), and therefore



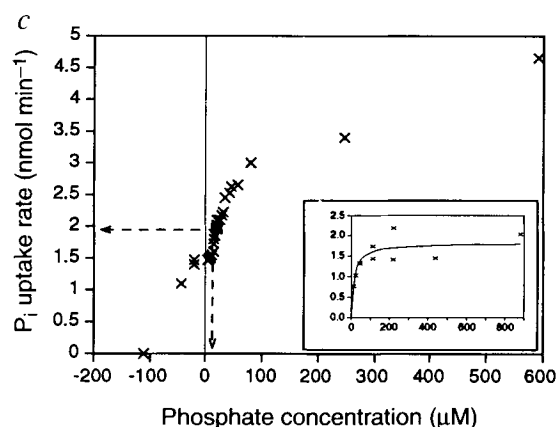
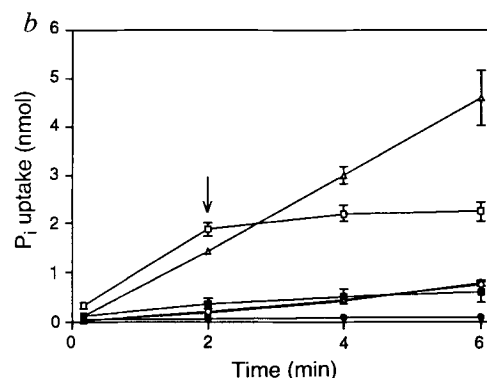
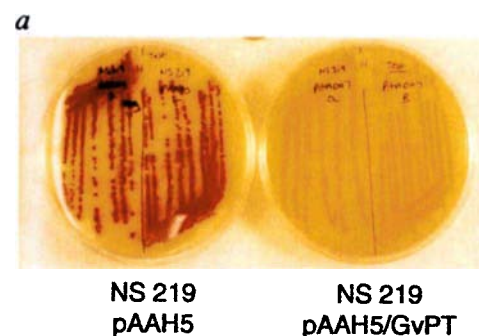
could be derived from the genome of either symbiont. Southern blot analysis with genomic DNA from *M. truncatula* suggested that GvPT did not represent a plant gene. As *G. versiforme* is an obligate symbiont, pure fungal material is limited. Therefore, to demonstrate that GvPT represented a gene from the *G. versiforme* genome, primers specific to the GvPT coding sequence were used in polymerase chain reaction (PCR) reactions to amplify the corresponding part of the GvPT sequence from gen-

omic DNA prepared from *G. versiforme* spores. The amplified fragment was sequenced to confirm its identity as part of the GvPT gene. In addition, PCR amplification of DNA prepared from mycorrhizal roots (*M. truncatula*/*G. versiforme*) resulted in the same DNA fragment, whereas amplification of DNA isolated from non-colonized *M. truncatula* roots was negative, as would be expected if GvPT represents a gene from the *G. versiforme* genome (data not shown).

|       |                                                    |     |
|-------|----------------------------------------------------|-----|
| GvPT  | MSTSDRVITDV-----DKRRAAL                            | 18  |
| PHO-5 | MSTPQK-----TAGGNAYHNFYNDLHFKDPNERRRLAL             |     |
| PHO84 | MSSVNDTIHVAERSLHKEHLTEGGNMAFNHNLNDFAHIEDPLERRRLAL  |     |
|       | ***                                                |     |
| GvPT  | KEIDDAKFGWQHIRACLVAGTGFFMDAYDLFAVNFAMSTIMGYVYGGKT- | 67  |
| PHO-5 | AEVDRAPFGWYHVRVAVAGVFFTFDSYDIFTVSLTLMLGIVYFQAKAR   |     |
| PHO84 | ESIDDEGFGWQVKTIISAGVGLTDSYDIFAINLGITMMSYVYWHGS--   |     |
|       | ***                                                |     |
| GvPT  | --PANIDLGLKVTGSGITLGLFFGYLADRLGRKRMYGVELMIIIVAT-   | 114 |
| PHO-5 | CLQPPTQP-----SSSQHRPARSLGRSASAPLL-----MSLVVR-ACTD  |     |
| PHO84 | -MPGFSQTLKLVSTSVGTIVIGQGFGLTADIVGRKRIYGMELIIMIVCT- |     |
|       | ***                                                |     |
| GvPT  | -----VASALSGESRAVTVVGTIMFWRVIMGVGIGGDY--           | 147 |
| PHO-5 | WNCSSSLPLSLRPWLRVTFHQHHR-----YHYLLACSYGGYRW        |     |
| PHO84 | -----ILQTTVAHSPAINFVAHSPAINFVAVLTFYRIVMGIGIGGDY--  |     |
|       | ***                                                |     |
| GvPT  | PLSAI--ITSEFATKKRRGMMASVFAMQGFILGSAIVALAVLAGFRNE   | 195 |
| PHO-5 | RLSSFOYITSEFATTKWRGAMGAVFAMQGLQLAAAFVTLGFKKS       |     |
| PHO84 | PLSSI--ITSEFATTKWRGAMGAVFAMQAGQISGGIIALILVAAYKGE   |     |
|       | ***                                                |     |
| GvPT  | IIKDVSAVDY---C-----WRIVLGGGAVPGLAALYFRLTIPETPRYT   | 235 |
| PHO-5 | LEAAPTLASCTGDCAVAVDKMWRTIVIGVAVPGCIALYRLTIPETPRYT  |     |
| PHO84 | LEYANSGAECDCARQKACDQWRILIGLTVLGLACLYFRLTIPESPRYQ   |     |
|       | ***                                                |     |
| GvPT  | MDVEHDVNKATSDITSYQKNDVNEDDNT-----GNHVGV-----P      | 272 |
| PHO-5 | FDVKRDVEQASDDIEAFKTKPKGQDEATRIVAKQEAKEMEI-----P    |     |
| PHO84 | LDVNAKLELAQAQEQDGEKKIHDTSDDEDMATNGLERASTAVESLDNHP  |     |
|       | ***                                                |     |
| GvPT  | KASWSDVSYFGKWSNGKVLGTSMSWFDIAFYGIGLNNGIILSAIGY     | 322 |
| PHO-5 | KASWGDFFRHYSKRNAMLAGTALSWCFDIAYYGVSLNNATILNVIGY    |     |
| PHO84 | KASFKDFCRHFGWKYKILLGTAGYWFLLDVAFYGLSLNSAVILQITIGY  |     |
|       | ***                                                |     |
| GvPT  | SETHADLNLRAYNSLKNMAYGNIIITIMGTAPGYVWTAALVDSWGRKPI  | 372 |
| PHO-5 | STTGAKN---TYEILYNTAVGNLIIVLGAVPGYVWTVFTVDTVGRKPI   |     |
| PHO84 | A--GSKN---VYKKLYDTAVGNLILCAGSLPGYVWVFTVDIIGRKPI    |     |
|       | ***                                                |     |
| GvPT  | QLMGFGVLTALFTVMGAAPNPLKEHSIPAFIILFTLLQFFQNFPGNTTTF | 422 |
| PHO-5 | QFMGFGILTILFVVMGFAYKHLSPH---ALLAIFVLQAQFFNFPGNATTF |     |
| PHO84 | QLAGFIILTALFCVIGFAYHKLGDH---GLLALYVICQFFQNFPGNTTTF |     |
|       | ***                                                |     |
| GvPT  | IVPGVFPTRYRSTGHGISAASGKLGAIYAQVGFSLKLDIGGSNA-----  | 467 |
| PHO-5 | IVPGVFPTRYRSTHGLSAAMKIGSIIQGGAIAPLRTRGAVKGGNPNP    |     |
| PHO84 | IVPGCFPTRYRSTAHGISAASGKVGAIQAQALGTLDHNCARDKPKTN    |     |
|       | ***                                                |     |
| GvPT  | -FVGPLLLIFSAMFFIGGLFSILIFETKGLSLEELN-----EEHTYDV   | 510 |
| PHO-5 | -WMNHVLEIYALFMLLVGTTFLIPETKRKTLEELSGEFDMSGEAAQRD   |     |
| PHO84 | CWLPHVMEIFALFMLLGIPTLLIPETKRKTLEEINELYHDEIDPATLNF  |     |
|       | ***                                                |     |
| GvPT  | EERKERIKA-----DA                                   | 521 |
| PHO-5 | TTLTEHKTEAPTSSAAVNA                                |     |
| PHO84 | RKNNDIESSSPQLQHEA                                  |     |

FIG. 1 Alignment of the deduced amino-acid sequence of GvPT with amino-acid sequences of yeast *PHO84* (ref. 9) and *Neurospora PHO-5* (ref. 20). Asterisks indicate conserved amino acids; bold letters indicate the predicted membrane-spanning domains of the deduced GvPT protein; boxed sequences are consensus sites for phosphorylation by protein kinase C; and boxed and shaded sequences are sites for phosphorylation by casein kinase II.

FIG. 2 Complementation of a yeast phosphate uptake mutant. **a**, Acid phosphatase activity in yeast *pho84* mutant NS219 containing either the vector (pAAH5) or the vector containing GvPT (pAAH5/GvPT) (two independent transformants for each construct) after growth on high-phosphate medium. Acid phosphatase activity is detected by staining. NS219 carrying the pAAH5 vector only shows *pho84* mutant phenotype and colonies stain red. NS219 carrying the vector expressing GvPT (pAAH5/GvPT) shows wild-type phenotype, and colonies remain pale. **b**, Uptake of phosphate by yeast *pho84* mutant NS219 containing either pAAH5 or pAAH5/GvPT. Phosphate uptake NS219, pAAH5/GvPT (open triangles), NS219 pAAH5 (filled triangles). Addition of CCCP at 2 mins (arrowed) for NS219, pAAH5/GvPT (open squares), NS219, pAAH5 (filled squares). Uptake in the absence of glucose: NS219, pAAH5/GvPT (open circles), NS219, pAAH5 (filled circles). Values are the averages from two independent transformants. Uptake values are nmol phosphate ml<sup>-1</sup> in yeast cell suspension with 1 absorbance unit at 600 nm. **c**, Direct linear plot of kinetic data to enable estimation of  $K_m$  and  $V_{max}$ <sup>19,20</sup>. Graph shows intersection points (lines omitted for clarity), arrows from the



median point (based on calculations<sup>19,20</sup>) show estimations of  $K_m$  (18  $\mu\text{M}$ ) and  $V_{max}$  (1.96 nmol min<sup>-1</sup> per absorbance unit of cells). Inset graph shows phosphate uptake rate in NS219 cells expressing GvPT as a function of phosphate concentration.

**METHODS.** The GvPT coding sequence was ligated into the yeast expression vector pAAH5 (ref. 16) to create pAAH5/GvPT and used to transform *S. cerevisiae* NS219 (*pho84* mutant)<sup>9</sup>. The pAAH5 plasmid (without insert) was also transformed into *S. cerevisiae* NS219 cells to act as a control. Growth of yeast strains, acid phosphatase overlay tests and phosphate uptake tests were performed by standard methods<sup>9,17,27</sup>.

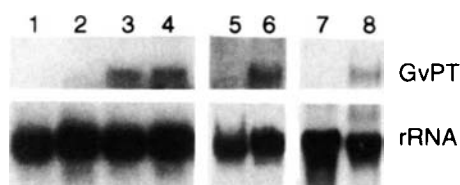


FIG. 3 Expression of GvPT in mycorrhizal roots. Northern blot of RNA from non-colonized roots of *M. truncatula* (two independent samples) (lanes 1 and 2), *M. sativa* (lane 5), *A. porrum* (lane 7) and RNA from colonized roots of *M. truncatula*/*G. versiforme* (two independent samples) (lanes 3 and 4), *M. sativa*/*G. versiforme* (lane 6), *A. porrum*/*G. versiforme* (lane 8). Plant growth and colonization procedures, RNA extraction, and northern blot analyses have been previously described<sup>28</sup>. The blots were hybridized with <sup>32</sup>P-labelled GvPT cDNA and with <sup>32</sup>P-labelled pSR1-2B3 (18S rRNA)<sup>29</sup>. The rRNA probe was used as a control for loading and transfer.

To determine whether GvPT encoded a functional phosphate transporter, the cDNA was ligated into a yeast expression vector<sup>16</sup> and transformed into a yeast strain NS219 carrying the *pho84* mutation<sup>9</sup>. Yeast *pho84* mutants lack the high-affinity phosphate transporter, and consequently have severely impaired phosphate uptake. Because they cannot accumulate phosphate, *pho84* mutant cells produce a repressible acid phosphatase (rAPase; EC 3.1.3.2), even during growth on high-phosphate media containing levels of phosphate sufficient to repress rAPase production in wild-type cells. Acid phosphatase production in yeast cells is easily detected by a simple staining technique<sup>17</sup>; *pho84* mutant cells stain red, while wild-type cells remain pale. NS219 transformants expressing GvPT displayed wild-type rAPase activity and remained pale, indicating that expression of GvPT complements the *pho84* mutant phenotype (Fig. 2a, right). NS219 transformants carrying the expression vector without the

GvPT sequence displayed the *pho84* phenotype and stained dark red (Fig. 2a, left). Phosphate uptake by NS219 cells expressing GvPT was further confirmed by measuring uptake of <sup>33</sup>P<sub>i</sub> from solution. NS219 expressing GvPT grew more rapidly (data not shown), and accumulated <sup>33</sup>P<sub>i</sub> at a rate significantly higher than the background rate shown by control NS219 cells (Fig. 2b). Uptake of <sup>33</sup>P<sub>i</sub> was decreased almost to background levels when glucose was omitted from the uptake media, indicating that energy was required for uptake. Addition of carbonyl-cyanide-*m*-chlorophenyl-hydrazone (CCCP), which uncouples the proton gradient across the membrane, prevented further <sup>33</sup>P<sub>i</sub> uptake, suggesting that the transporter is operating by proton-coupled symport (Fig. 2b). This transport mechanism has been demonstrated previously for the yeast *PHO84* transporter<sup>18</sup>. Phosphate uptake by NS219 cells expressing GvPT follows Michaelis-Menten kinetics with an apparent *K<sub>m</sub>* of 18 μM<sup>19,20</sup> (Fig. 2c). This is similar to the yeast high-affinity phosphate transport system for which the *K<sub>m</sub>* has been estimated to be 8.2 μM in whole yeast cells<sup>9</sup>, and 24 μM for the PHO84 protein expressed in proteoliposomes<sup>18</sup>.

Previous measurements of phosphate uptake in hyphae from germinating spores of the arbuscular mycorrhizal fungus *Gigaspora margarita* suggested that this mycorrhizal fungus, like other fungi<sup>21</sup>, probably possesses two systems for phosphate uptake: an active, high-affinity system (*K<sub>m</sub>* 1.8–3.1 μM), and passive, low-affinity phosphate transport system (*K<sub>m</sub>* 10.2–11.3 mM (ref. 22)). The *K<sub>m</sub>* and transport properties of GvPT are consistent with those of a high-affinity uptake system.

GvPT is expressed during mycorrhizal associations with the leguminous hosts *M. truncatula* and alfalfa (*M. sativa*), as well as a non-legume, leek (*Allium porrum*) (Fig. 3). *In situ* hybridization experiments indicated that GvPT was not expressed in fungal structures within the root (data not shown). Therefore, by implication, GvPT is probably expressed in the external hyphae that ramify out of the root into the soil. To determine unequivocally

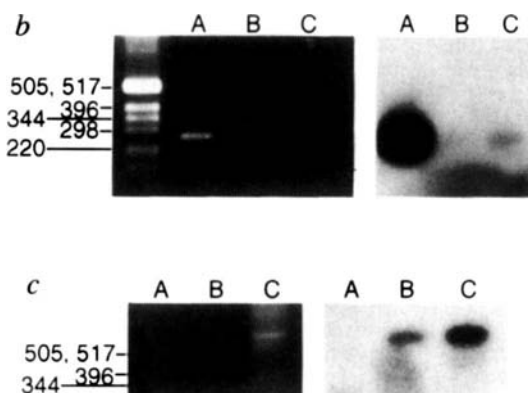
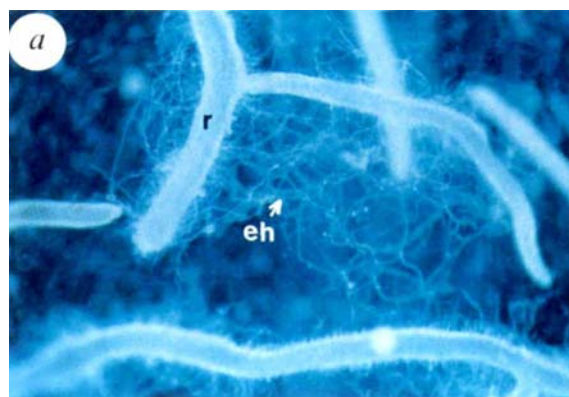


FIG. 4 Localization of expression of GvPT in external hyphae extending from mycorrhizal roots. a, Mycorrhizal roots (r) showing external hyphae (eh). b, Agarose gel (left) and corresponding Southern blot (right) of RT-PCR products amplified using primers specific for GvPT coding sequence demonstrates that GvPT is expressed in the external hyphae. RNA used in the RT-PCR reactions was as follows: lane A, RNA from external hyphae of *G. versiforme*; lane B, as lane A except the RT-PCR reaction was performed without reverse transcriptase; lane C, RNA from a root piece containing internal hyphae and arbuscules of *G. versiforme* but lacking external hyphae. The PCR product was not obtained from non-colonized *M. truncatula* roots (data not shown). Southern blot (right) of agarose gel (left) was hybridized with <sup>32</sup>P-labelled GvPT cDNA. c, Agarose gel (left) and corresponding Southern blot (right) of RT-PCR products amplified using primers specific for mycorrhizal fungal rRNA genes indicates the level of RNA in the external and internal samples and allows quantification of relative fungal RNA levels. RNA used in the RT-PCR reactions was as follows: lane A, RNA from non-colonized root piece; lane B, RNA

from a root piece containing internal hyphae and arbuscules of *G. versiforme* but lacking external hyphae; lane C, RNA from external hyphae from *G. versiforme*. DNA markers (sizes in bp) are shown to the left of the figures. Southern blot (right) of agarose gel (left) was hybridized with <sup>32</sup>P-labelled rDNA sequence.

**METHODS.** External hyphae projecting from the surface of mycorrhizal roots of *A. porrum* were visualized under a binocular microscope and extracted with forceps. About 1 mg of hyphae was collected per sample. Densely colonized root pieces were selected similarly and any external hyphae protruding from the root were removed with forceps or brushed off with a fine paintbrush. Short root segments of 5–8 mm were collected. Half of these pieces were placed in ethanol and stained to confirm colonization<sup>28</sup>. RNA extraction and RT-PCR analysis were performed by standard procedures<sup>28,30</sup>. GvPT primers had the following sequences: (1) 5'-TGGTCATGGTATATCAGCTG-3'; (2) 5'-ATGCGTCAGCCTTTATCC-3'. rRNA primers were VANS1 and NS21 (ref. 23). Phosphor-image analysis of the Southern blots enabled quantification of the transcript levels.

the precise location of GvPT expression, external hyphae were collected from intact mycorrhizal roots (Fig. 4a), and the presence of GvPT transcripts was subsequently analysed by reverse transcription (RT)-PCR. Samples of root tissue containing internal, but not external, hyphae were also sampled and analysed to determine whether GvPT could be detected in the fungal structures within the root. These analyses indicated the presence of GvPT transcripts in external hyphae, and at extremely low levels in internal fungal structures within the root (Fig. 4b). The amount of fungal RNA, as estimated using rRNA-specific primers<sup>23</sup>, was slightly higher (1.83-fold) in the hyphal sample than the root sample (Fig. 4c). However, GvPT transcripts were 30-fold higher in the external hyphae, indicating that this is the main site of expression of this gene (Fig. 4b,c). The function and site of expression of GvPT are consistent with a role in the initial uptake of phosphate from the soil.

The mechanisms underlying phosphate transfer from the fungus to the plant remain unknown, but we speculate that the process occurs at the arbuscular interface, and potentially requires two transmembrane transport steps: the first to enable efflux of phosphate from the fungus, and the second to mediate uptake of phosphate by the plant<sup>24</sup>. The nature of these carriers remains to be described.

Cloning and initial characterization of GvPT has provided the first information about the molecular mechanisms underlying phosphate transport by the fungal partner in a VA mycorrhizal symbiosis. This will enable detailed investigations of the mechanisms regulating phosphate uptake in the mycorrhizal symbiosis, and will further our understanding of the differences in phosphate transport efficiency observed in different mycorrhizal fungi<sup>25,26</sup> and the mechanisms of phosphate acquisition by higher plants. □

Received 30 August; accepted 10 October 1995.

1. Harley, J. L. & Smith, S. E. *Mycorrhizal Symbiosis* (Academic, London, 1983).
2. Jeffries, P. *Crit. Rev. Biotechnol.* **5**, 319–357 (1987).
3. Smith, S. E. & Gianinazzi-Pearson, V. A. *Rev. Pl. Physiol. Pl. molec. Biol.* **39**, 221–244 (1988).
4. Sanders, F. E. & Tinker, P. B. *Nature* **233**, 278–279 (1971).
5. Pearson, J. N. & Jakobsen, I. *New Phytol.* **124**, 489–494 (1993).
6. Smith, S. E. *Biol. Rev.* **55**, 475–510 (1980).
7. Bethlenfalvay, G. W. *Am. Soc. Agron. Spec. Publ.* **54**, 1–27 (1992).
8. Pfleger, F. L. & Linderman, R. G. *Mycorrhizae and Plant Health* (APS symposium series, St Paul, MN, 1994).
9. Bun-ya, M., Nishimura, M., Harashima, S. & Oshima, Y. *Molec. cell. Biol.* **11**, 3229–3238 (1991).
10. Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. J. *molec. Biol.* **179**, 125–142 (1984).
11. Griffith, J. K., Baker, M. E., Duncan, R. A., Page, M. G. P. & Skurra, R. A. *Curr. Opin. Cell Biol.* **4**, 684–695 (1992).
12. Marger, M. D. & Saier, M. H. Jr. *Trends biochem. Sci.* **18**, 13–20 (1993).
13. Mann, B. J., Bowman, B. J., Grotelueschen, J. & Metzner, R. L. *Gene* **83**, 281–289 (1989).
14. Versaw, W. K. *Gene* **153**, 135–139 (1995).
15. Lawrence, J. C. Jr, Hiken, J. F. & James, D. E. *J. biol. Chem.* **265**, 2324–2332 (1990).

16. Ammerer, G. *Meth. Enzym.* **101**, 192–201 (1983).
17. Toh-E, A. & Oshima, Y. *J. Bact.* **120**, 608–617 (1974).
18. Berhe, A., Fristedt, U. & Persson, B. L. *Eur. J. Biochem.* **227**, 566–572 (1995).
19. Brady, J. F. & Ishizaki, H. *Comput. Meth. Program. Biomed.* **28**, 271–272 (1989).
20. Eisenberg, R. & Cornish-Bowden, A. *Biochem. J.* **139**, 715–720 (1974).
21. Beever, R. E. & Burns, D. J. W. *Adv. bot. Res.* **8**, 127–219 (1980).
22. Thomson, B. D., Clarkson, D. T. & Brain, P. *New Phytol.* **116**, 647–653 (1990).
23. Simon, L., Lalonde, M. & Bruns, T. D. *Appl. envir. Microbiol.* **58**, 291–295 (1992).
24. Smith, S. E. & Smith, F. A. *New Phytol.* **114**, 1–38 (1990).
25. Jakobsen, I., Abbott, L. K. & Robson, A. D. *New Phytol.* **120**, 371–380 (1992).
26. Pearson, J. N. & Jakobsen, I. *New Phytol.* **124**, 481–488 (1993).
27. Ueda, Y. & Oshima, Y. *Molec. gen. Genet.* **136**, 255–259 (1975).
28. Harrison, M. J. & Dixon, R. A. *Pl. J.* **6**, 9–20 (1994).
29. Eckenrode, V. K., Arnold, J. & Meagher, R. B. *J. molec. Evol.* **21**, 259–269 (1985).
30. Kawasaki, E. S. in *PCR Protocols* (eds Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J.) 21–28 (Academic, San Diego, 1990).

ACKNOWLEDGEMENTS. We thank Y. Oshima and S. Harashima for PMB13 and NS219; R. A. Dixon and R. S. Nelson for critical reviews of the manuscript; M. Guttenberger for assistance with the Brady and Ishizaki computer program; and A. Wilkins for assistance with the final preparation of the manuscript. This work was supported by The Samuel Roberts Noble Foundation.

## A channel-like transporter for $\text{NH}_4^+$ on the symbiotic interface of $\text{N}_2$ -fixing plants

Stephen D. Tyerman, Lynne F. Whitehead\* & David A. Day\*

Adelaide Centre for Plant Membrane Biology, School of Biological Sciences, The Flinders University of South Australia, GPO Box 2100, Adelaide 5001, Australia

\* Division of Biochemistry and Molecular Biology, Faculty of Science, Australian National University, ACT 0200, Australia

**SYMBIOSIS with nitrogen-fixing bacteria (rhizobia) allows legumes to survive in nitrogen-poor soils. The nitrogen-fixing bacteroids are found inside root nodule cells within the symbiosome, an organelle bounded by the peribacteroid membrane<sup>1</sup>. Across this membrane the plant receives fixed nitrogen in exchange for reduced carbon<sup>2</sup>. It has been assumed that fixed nitrogen is released from the symbiosome as either  $\text{NH}_3$  or  $\text{NH}_4^+$ , but until now the transport mechanism was unknown<sup>3–5</sup>. We report here the use of patch-clamp techniques to show, in membrane patches on isolated symbiosomes, smoothly activating currents that are passive and equivalent to the influx of cations to the plant cytoplasm. The currents are largest with  $\text{NH}_4^+$  as the cation at physiological concentration and they are blocked by calcium ions on the bacteroid side of the membrane. The characteristics of the currents are more like those of a channel than of carrier-mediated transport. *In vivo* nitrogen would move passively as  $\text{NH}_4^+$  to the cytoplasm upon energization of the membrane and calcium ions may be involved in regulation of the flux.**

Soya bean/*Bradyrhizobium japonicum* symbiosomes are 2–5  $\mu\text{m}$  in diameter and enclose 2–10 bacteroids<sup>6</sup>. Using a patch-

clamp with small pipette tips, gigaohm seals were formed with the peribacteroid membrane (PBM) (Fig. 1). In all inside-out patches ( $n=48$ ), time-dependent inward currents (into pipette) occurred when univalent cations were present in the bath and when the membrane potential was made more negative (cytoplasm with respect to symbiosome) (Fig. 1a). Increasing  $\text{NH}_4\text{Cl}$  in the bath increased the inward currents (Fig. 1a–c) and shifted the reversal potential of the current to positive (Fig. 1d). Therefore the current was carried by the flux of  $\text{NH}_4^+$  into the pipette, equivalent to  $\text{NH}_4^+$  leaving the symbiosome. Current–voltage curves from fast ramps on the inward current (Fig. 1d) showed that the reversal potentials behaved as simple cation diffusion potentials with equal permeability to  $\text{NH}_4^+$  and  $\text{K}^+$ . The ratios of the current at  $-100\text{ mV}$  for test ions relative to that for  $\text{NH}_4^+$  show that the transport system is non-selective between the physiologically important ions ( $\text{NH}_4^+$ ,  $\text{K}^+$ ,  $\text{Na}^+$ ) at high concentration (150 mM) but gave larger currents for  $\text{NH}_4^+$  relative to  $\text{K}^+$  at lower concentrations (20 mM, Fig. 1e). Divalent cations and choline did not give measurable inward currents. The conductance of patches displayed two classes of concentration dependence (Fig. 1f). For the saturating class the apparent binding constant ( $K_m$ ) was 37.5 mM (95% conf.: 20.5–54.5 mM), close to the estimated values for  $\text{NH}_4^+$  within the symbiosome<sup>7</sup>. An explanation for this observation is that at the time of isolation all the  $\text{NH}_4^+$  transporters were set in one of two affinity states. Other measured properties of the inward current were the same for the two classes. There was no coupling with a proton gradient because the reversal potential did not shift for pHs between 5 and 8. These observations indicate that the transport system accounting for the  $\text{NH}_4^+$  currents is a passive channel-mediated conductance.

The  $\text{Ca}^{2+}$  content of symbiosomes is reported to be high (A. H. Millar and D.A.D., unpublished results). Calcium ions in the bath blocked the inward current (Fig. 2a, b). At  $\text{NH}_4^+$  concentra-

FIG. 1 Inward currents carried by  $\text{NH}_4^+$  across inside-out patches from symbiosomes. **a**, **b**, Current-time curves in response to voltage pulses for two  $\text{NH}_4^+$  concentrations. **c**, Current-voltage curves of time-dependent current in a range of  $\text{NH}_4^+$  concentrations. **d**, Difference  $I/V$  curves from ramps<sup>27</sup> that capture the true reversal potential of the current. The reversal potentials shifted in a nernstian manner ( $-0.3$ ,  $-26.7$ ,  $-60.4$ ,  $-101.5$  mV) with changes in  $\text{NH}_4\text{Cl}$  concentration in the bath (given on graph). Currents filtered at 1 kHz, digitized at 5 kHz. **e**, Current of test ion (at 150 mM, stippled; or 20 mM,  $\text{K}^+$  only) relative to that for  $\text{NH}_4^+$  at  $-100$  mV. Values are mean ( $\pm$  s.e.m.,  $n=3$ ) except for  $\text{K}^+$  ( $n=4$ ) and for  $\text{Na}^+$  ( $n=2$ ); met, methylamine; cho, choline. **f**,  $\text{NH}_4^+$  concentration dependence of conductance at  $-120$  mV (normalizing to 30 mM  $\text{NH}_4^+$ ). A linear relationship was observed in four patches. The Michaelis-Menten equation was fitted using a non-linear least-squares method ( $n=5$ ). Error bars are  $\pm$  s.e.m. METHODS. Symbiosomes (3 to 5  $\mu\text{m}$  diameter) were isolated<sup>6</sup> from *Glycine max* (L. cv. Stephens or Forest) inoculated with *Bradyrhizobium japonicum* USDA 110. Symbiosomes remained intact in an initial bathing solution of (in mM): 100 potassium glutamate, 2  $\text{MgCl}_2$ , 2.3  $\text{CaCl}_2$ , 10 EGTA, 10 HEPES/Tris pH 7. Symbiosomes adhered to the base of the chamber<sup>27</sup> and could be viewed at  $\times 800$  magnification. Patch pipettes<sup>27</sup> (40 to 200 M $\Omega$ ) yielded seals (typically 100 Gohm) when filled with (in mM): 150 KCl, 10  $\text{CaCl}_2$ , 5 HEPES, pH 7.2 (KOH). Only inside-out patches with no vesicle or adhering bacteroids were used. Transitions of occasional large channels indicated that a vesicle was not attached<sup>28</sup>. Currents were measured and corrected<sup>28</sup> as described previously<sup>29</sup>. Unless indicated otherwise current output was filtered at 100 Hz and digitized at 200 Hz. All test bath solutions consisted of (in mM), 5 HEPES/Tris pH 7, test cation as Cl salt. Solutions were made to 400 mosmol  $\text{kg}^{-1}$  with mannitol.

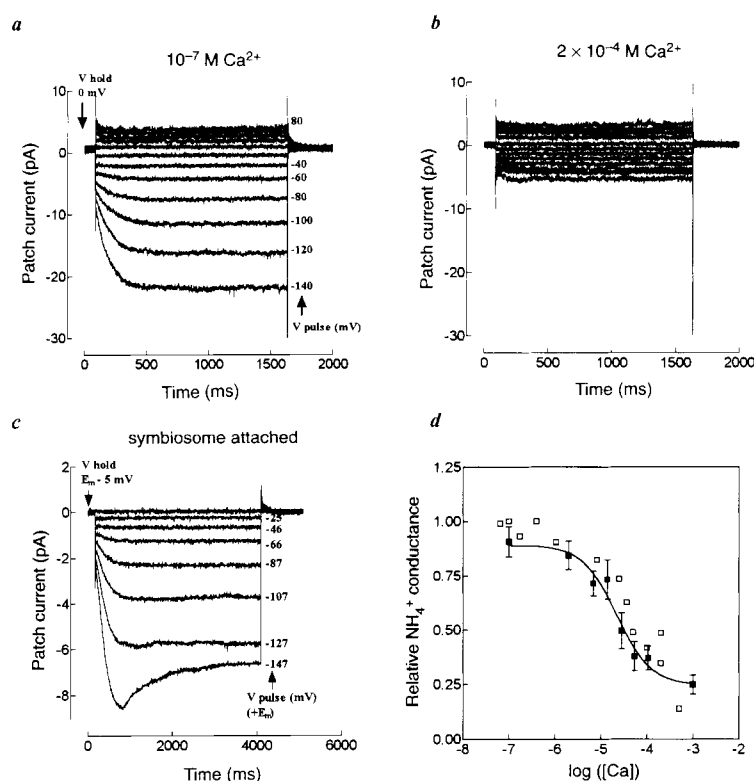
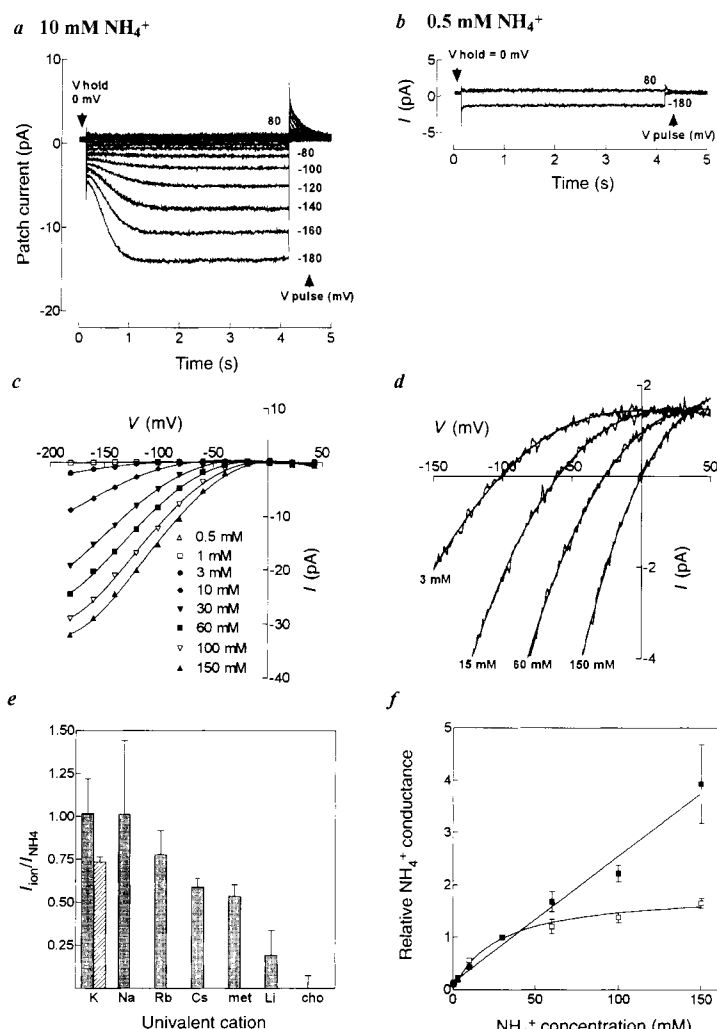
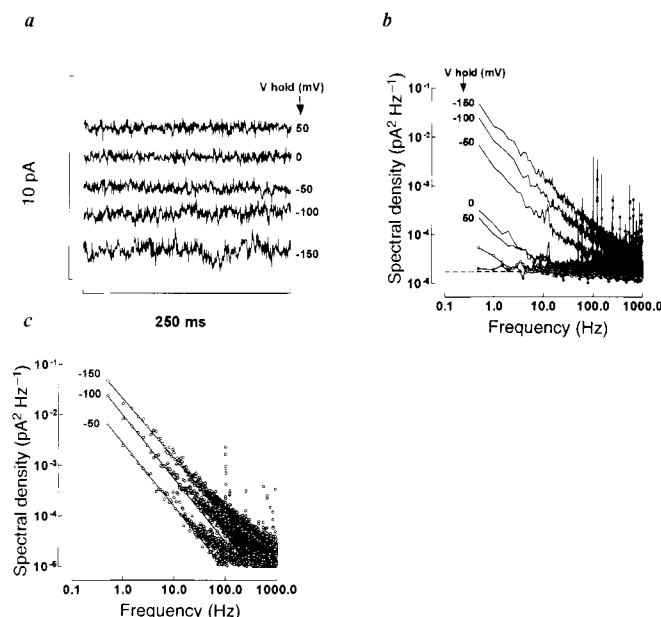


FIG. 2 Effect of  $\text{Ca}^{2+}$  on the time-dependent inward current and comparison with a symbiosome attached patch. **a**, **b**, Current versus time curves for an inside-out patch bathed in 15 mM  $\text{NH}_4^+$  with two bath  $\text{Ca}^{2+}$  concentrations. The pipette contained initial bath solution ( $\sim 130$  nM free  $\text{Ca}^{2+}$ ). The time-dependent currents were similar to those normally seen with 10 mM  $\text{Ca}^{2+}$  in the pipette. **c**, Current-time curves for a symbiosome-attached patch illustrating that the time-dependent current is present despite reported high total  $\text{Ca}^{2+}$  concentration in symbiosomes. The inactivation may be due to depletion of univalent cation in the symbiosome space<sup>28</sup>. The bath contained initial bath solution. **d**, Plot of relative  $\text{NH}_4^+$  conductance at  $-120$  mV (error bars  $\pm$  s.e.m.,  $n=5$ ) versus bath free  $\text{Ca}^{2+}$  concentration where the pipette contained 10 mM  $\text{Ca}^{2+}$  (filled symbols). The bath contained 20 mM  $\text{NH}_4^+$  and  $\text{Ca}^{2+}$  was buffered with 2 mM EGTA/Tris. Also shown are the results from two patches (open symbols) where the pipette  $\text{Ca}^{2+}$  was buffered at 130 nM (see **a**, **b**). The line is a least-squares fit of a dose-response curve with variable Hill slope to the high pipette  $\text{Ca}^{2+}$  data.



tions close to the apparent  $K_m$ , the inhibitory constant ( $K_i$ ) for  $\text{Ca}^{2+}$  was  $22 \mu\text{M}$  (95% conf.:  $9.9\text{--}49.5 \mu\text{M}$ ,  $n=5$ ) and Hill slope = 1.2 (95% conf.:  $0.2\text{--}2.1$ ) (Fig. 2d). The characteristics of the current were independent of  $\text{Ca}^{2+}$  concentration in the pipette (Fig. 2). In symbiosome-attached patches, time-dependent currents were also observed and in some cases showed inactivation (Fig. 2c). The currents were generally small due to either a low univalent cation concentration in the symbiosome (compare with Fig. 1a) or a blocking  $\text{Ca}^{2+}$  concentration. The symbiosome-attached data suggests that the high total  $\text{Ca}^{2+}$  measured in symbiosomes is mostly bound in the bacteroids<sup>8</sup> and that the free  $\text{Ca}^{2+}$  concentration in the symbiosome space is in a range that allows  $\text{NH}_4^+$  transport.

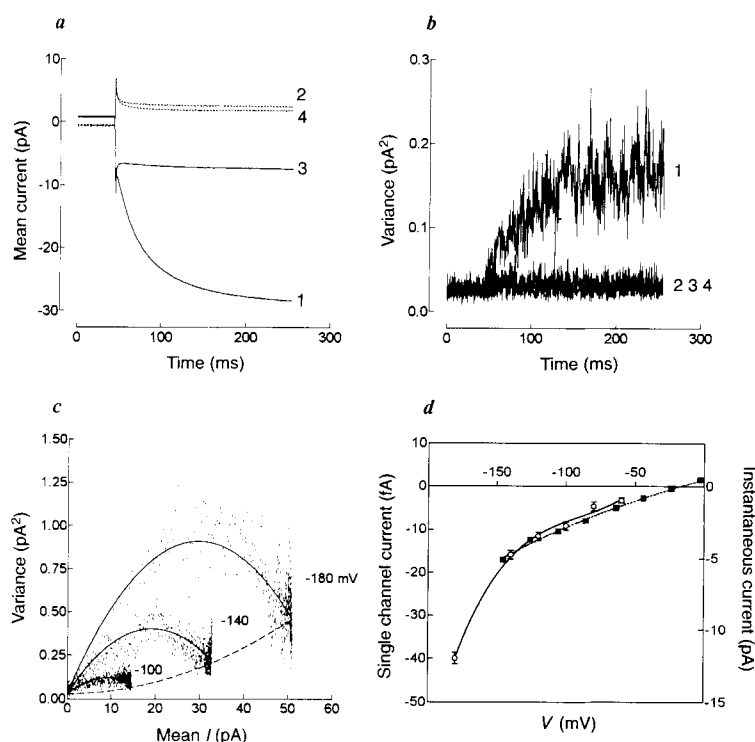


The passive nature of the current indicates that ion channels may be responsible; however, because no discrete single-channel openings were detected, we tested whether the current noise had channel characteristics. Concomitant with an increase in inward current there was an increase in noise (Figs 3a and 4b). The noise spectrum showed a strong voltage and frequency dependence between 0.4 and 1,000 Hz and was well above that associated with the patch-clamp amplifier and pipette (Fig. 3b). Subtracting the spectra for 50 mV, as the background noise, from the spectra at negative potentials gives essentially  $1/f$  type spectra (Fig. 3c). This has been observed for a number of ion channels<sup>9,11</sup> and may indicate either cooperativity<sup>10</sup> or a channel with at least four kinetic states, the minimum number that could account

FIG. 3 Noise associated with the time-dependent inward current in a detached patch bathed in 150 mM  $\text{NH}_4^+$ . a, Examples of current traces at different membrane potentials illustrating that an increase in noise can be observed in the current at membrane potentials at which the inward current is activated. b, Noise spectra of patch current at different membrane potentials and the background noise associated with the amplifier (filled circles, head stage) and a pipette pushed against a bead of sylgard (open symbols). The dashed line is a theoretical estimate of the amplifier noise and quantization noise during digitization, and was largely accounted for by thermal noise in the feedback resistor. The variances for each of the records used for the spectra are: head stage,  $0.020 \text{ pA}^2$ ; pipette against sylgard,  $0.024 \text{ pA}^2$ ; PBM inside-out patch at: 50 mV,  $0.031 \text{ pA}^2$ ; 0 mV,  $0.032 \text{ pA}^2$ ; -50 mV,  $0.068 \text{ pA}^2$ ; -100 mV,  $0.129 \text{ pA}^2$ ; -150 mV,  $0.216 \text{ pA}^2$ . c, Difference power spectra after subtraction of the spectra for 50 mV point by point. The fitted lines are to the equation  $S(f) = af^{-k}$  where  $k$  equalled 1.2.

METHODS. For noise analysis (Figs 3 and 4) data was sampled at 2 kHz and filtered at 1 kHz (4 pole Bessel). The noise spectra were generated from records of 131,072 points, (about 65 s). After digitally removing linear DC components the spectra were generated using the Welch periodogram technique with 64 segments and 50% overlap. For non-stationary noise analysis sampling and filtering at higher rates did not give significant increases in the single-channel current<sup>16</sup>. The 14 bit A/D converter ( $\pm 8 \text{ V}$ ) and amplifier gain used ( $100 \text{ mV pA}^{-1}$ ) ensured that the quantization error did not cause a significant underestimation of the single-channel current<sup>14</sup>.

FIG. 4 Non-stationary noise analysis on the inward current. a, Average currents from 40 consecutive pulses +40 mV to -120 mV (curves 1 and 3) and 0 mV to 100 mV (curves 2 and 4). The bath contained 150 mM KCl (curves 1 and 2) or 150 mM KCl plus 10 mM  $\text{CaCl}_2$  (curves 3 and 4).  $\text{Ca}^{2+}$  inhibited the current and abolished the increase in variance (b, curve 3). A positive pulse in the presence of  $\text{Ca}^{2+}$  indicated that no significant change in leak current had occurred (curve 4). b, Variance of the currents shown in a calculated according to ref. 15. The mean/variance plot for this example yielded a single-channel amplitude ( $i$ ) of 7 fA at -120 mV. c, Variance/mean plots from data obtained as in a and b on a different patch bathed in 150 mM  $\text{NH}_4^+$  at three membrane potentials (-100, -140, -180 mV). The leak current has been subtracted from the mean current ( $I$ ). The fitted curves are given by the equation,  $\sigma^2 = iI - I^2/N$ . The number of channels ( $N$ ) was initially allowed to vary for each fit.  $N$  (close to 1,000) for each potential above -100 mV was then held constant for the fits shown. The baseline variance was  $0.03 \text{ pA}^2$ . The dashed line is given by  $\sigma^2 = kI^2$  ( $k = 1.65 \times 10^{-4}$ ) as a maximum estimate of transport noise to provide a lower limit on  $N$ . The patch area in this case was estimated to be between  $0.4$  to  $3 \mu\text{m}^2$  (different possible geometries<sup>30</sup>), from the pipette resistance and the measured radius of similar pipettes. (d) Single-channel current versus voltage (open symbols,  $\pm$  s.e. from fit) and instantaneous current from a patch in identical solutions obtained from tail currents (filled symbols) when the membrane potential was held at -150 mV and pulsed positive. The axes are scaled so that the single-channel curve and instantaneous curves correspond illustrating that over the same voltage range the currents are directly proportional.



for the data in Fig. 3c with 90% confidence limits<sup>11</sup>. The noise due to gating of the channel may also be swamped by transport noise through predominantly open channels if  $P_{\text{open}}$  is high due to a steep activation with voltage. Boltzmann plots support this conclusion (also see Fig. 4c). Because carrier transport noise has been suggested to increase with increasing frequency<sup>12</sup>, we conclude that the  $\text{NH}_4^+$  transporter is more like a channel confirming the conclusions drawn from the data in Fig. 1.

Lorentzian components could not be separated from the noise spectra so we used non-stationary noise analysis to estimate an equivalent single-channel conductance. This is not affected by the number of kinetic states but assumes that there are only two conductance states<sup>13</sup>. The analysis involves determining the current variance during a time-dependent change in probability of opening ( $P_{\text{open}}$ ) as occurs during activation by negative voltage pulses (Fig. 4a, b). Fitting of the data to the theoretical quadratic equation of variance as a function of mean current<sup>13–15</sup> at different membrane potentials (Fig. 4c) provides estimates of the single-channel current and number of channels. The single channel  $I/V$  curve (Fig. 4d) compared favourably with the instantaneous  $I/V$  curve. The conductance of the single channels measured in the linear region of the  $I/V$  curve was 0.11 pS (s.e. = 0.021,  $n = 3$  patches). The upper estimate of the number of channels in the patch for Fig. 4c was 1,000. A lower estimate was 840 if the final variance at  $-180$  mV was attributed to transport noise that increases with the square of current<sup>16</sup> (dashed line, Fig. 4c) and assuming that the  $P_{\text{open}}$  at full activation was one (indicated from Boltzmann plots). The range in channel density was between 300 and 2,400  $\mu\text{m}^{-2}$ . This density is high for plant membranes, but does not approach some of the estimates made for animal and bacterial membranes.

The average conductance of the ensemble  $\text{NH}_4^+$  channels in a patch was 37 pS (at 10 mM  $\text{NH}_4^+$ ). Estimates of the  $\text{NH}_4^+$  flux *in vivo* per symbiosome<sup>17</sup> converts to an equivalent current of 12 fA. This can be accommodated by the  $\text{NH}_4^+$  channel with only a 0.01 mV negative bias due to operation of the proton pump (using patch area to scale to a symbiosome conductance of 780 pS). Our data support the hypothesis that the PBM proton pump ( $\text{H}^+$ -ATPase)<sup>3</sup> acidifies the symbiosome space converting  $\text{NH}_3$  to  $\text{NH}_4^+$ , thereby creating an  $\text{NH}_3$  concentration gradient from the bacteroids. The high  $\text{NH}_4^+$  concentration in the symbiosome space and low concentration in the host cytoplasm<sup>7</sup>, together with the negative voltage generated by the pump, would drive  $\text{NH}_4^+$  through the inward rectifying  $\text{NH}_4^+$  channel.

The PBM is plant derived with similarities to the plasma membrane<sup>18</sup>, therefore the  $\text{NH}_4^+$  transport we have characterized could occur in the plasma membrane and may underlie other passive  $\text{NH}_4^+$  uniporters<sup>19</sup>. Sub-picoSeimen channels have been characterized in animal cell membranes<sup>20,21</sup> but have not, until now, been studied in detail in plants<sup>22,23</sup>. This  $\text{NH}_4^+$  channel shows little similarity to other transporters permeable to  $\text{NH}_4^+$  that have been characterized, including: the  $\text{K}^+$  inward rectifier channel from plant plasma membrane<sup>24</sup>, the low-affinity  $\text{NH}_4^+$  transporter from yeast which is not blocked by  $\text{Ca}^{2+}$  (ref. 25); and the high affinity  $\text{NH}_4^+$  transporter from *Arabidopsis*<sup>26</sup>. We suggest that the channel plays an important role in the delivery of fixed nitrogen from the bacteroid to the plant. □

13. Sigworth, F. J. *J. Physiol.* **307**, 97–129 (1980).
14. Heinemann, S. H. & Conti, F. *Meth. Enzym.* **207**, 131–148 (1992).
15. Silberberg, S. D. & Magleby, K. L. *Biophys. J.* **65**, 1570–1584 (1993).
16. Hainsworth, A. H., Lewis, R. A. & Eisenberg, R. S. *J. gen. Physiol.* **104**, 857–883 (1994).
17. Bergersen, F. J. *Protoplasma* **183**, 49–61 (1994).
18. Brewin, N. J. et al. *EMBO J.* **4**, 605–611 (1985).
19. Frommer, W. B. et al. *Pl. molec. Biol.* **26**, 1651–1670 (1994).
20. Bodoia, R. D. & Detwiler, P. B. *J. Physiol.* **367**, 183–216 (1984).
21. Fesenko, E. E., Kolesnikov, S. S. & Lyubarsky, A. L. *Nature* **313**, 310–313 (1985).
22. Tyerman, S. D. A. *Rev. Pl. Physiol. Pl. molec. Biol.* **43**, 351–373 (1992).
23. Hedrich, R. & Becker, D. *Pl. molec. Biology* **26**, 1637–1650 (1994).
24. Schroeder, J. I., Ward, J. M. & Gassmann, W. A. *Rev. Biophys. biomolec. Struct.* **23**, 441–471 (1994).
25. Marini, A.-M., Vissers, S., Urrestarazu, A. & André, B. *EMBO J.* **13**, 3456–3463 (1994).
26. Ninnemann, O. W., Jauniaux, J. C. & Frommer, W. B. *EMBO J.* **13**, 3464–3471 (1994).
27. Schachtman, D. P., Tyerman, S. D. & Terry, B. R. *Pl. Physiol.* **97**, 598–605 (1991).
28. Barry, P. H. & Lynch, J. W. J. *Membr. Biol.* **121**, 101–117 (1991).
29. Skerrett, M. & Tyerman, S. D. *Planta* **192**, 295–305 (1994).
30. Sakmann, B. & Neher, E. in *Single Channel Recording* (eds Sakmann, B. & Neher, E.) 37–51 (Plenum, New York, 1983).

ACKNOWLEDGEMENTS. The expert technical assistance of D. Trussell and the help of C. Niemietz, G. Findlay and K. Schuller (and laboratory) are gratefully acknowledged. Thanks to D. Schachtman for reading the manuscript. This work was funded by Australian Research Council grants to S.D.T. & D.D.

## Role of auxilin in uncoating clathrin-coated vesicles

Ernst Ungewickell, Huberta Ungewickell, Susanne E. H. Holstein, Robert Lindner, Kondury Prasad\*, Winifred Barouch\*, Brian Martin†, Lois E. Greene\* & Evan Eisenberg\*

Center for Immunology, Department of Pathology, Washington University School of Medicine, 660 South Euclid Avenue, St Louis, Missouri 63110, USA

\* Laboratory of Cell Biology, National Institutes of Health, Bethesda, Maryland 20892, USA

† Clinical Neuroscience Branch, National Institutes of Health, Bethesda, Maryland 20892, USA

CLATHRIN-coated vesicles transport selected integral membrane proteins from the cell surface and the *trans*-Golgi network to the endosomal system<sup>1,2</sup>. Before fusing with their target the vesicles must be stripped of their coats. This process is effected by the chaperone protein hsp70c together with a 100K cofactor<sup>3</sup> which we here identify as the coat protein auxilin. Auxilin binds with high affinity to assembled clathrin lattices and, in the presence of ATP, recruits hsp70c. Dissociation of the lattice does not depend as previously supposed on clathrin light chains or on the amino-terminal domain of the heavy chain<sup>4,5</sup>. The presence of a J-domain at its carboxy terminus now defines auxilin as a member of the DnaJ protein family. In conjunction with hsp70, DnaJ proteins catalyse protein folding, protein transport across membranes and the selective disruption of protein–protein interactions<sup>6–8</sup>. We show that deletion of the J-domain of auxilin results in the loss of cofactor activity.

The purification of the 100K cofactor required for hsp70c-dependent dissociation of clathrin baskets parallels that of the coat protein auxilin<sup>3,9</sup>. This suggested to us that the two proteins might be one and the same. We found that auxilin comigrated with the cofactor in sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE), reacted on immunoblots with the auxilin-specific monoclonal antibody 100/4 (ref. 9) and the sequences of 9 peptides obtained by microsequencing of the 100K cofactor were identical to those of the following segments of cow brain auxilin: 71–81, 351–361, 507–518, 572–583 (1 unidentified residue), 613–622, 651–660, 811–830 (2 unidentified residues), 819–826 and 889–900. To determine whether auxilin possesses cofactor activity, the protein was removed from a brain coat protein extract by adsorption on an antibody 100/4 matrix<sup>10</sup>. Clathrin and associated proteins were reassembled into

Received 14 April; accepted 10 October 1995.

1. Day, D. A. & Udvardi, M. K. *Symbiosis* **14**, 175–189 (1992).
2. Udvardi, M. K., Price, G. D., Gresshoff, P. M. & Day, D. A. *FEBS Lett.* **231**, 36–40 (1988).
3. Blumwald, E., Fortin, M. G., Rea, P. A., Verma, D. P. S. & Poole, R. J. *Pl. Physiol.* **78**, 665–672 (1985).
4. Shearer, G. & Kohl, D. H. *Aust. J. Pl. Physiol.* **16**, 305–313 (1989).
5. Udvardi, M. K. & Day, D. A. *Pl. Physiol.* **94**, 71–76 (1990).
6. Day, D. A., Price, G. D. & Udvardi, M. K. *Aust. J. Pl. Physiol.* **16**, 69–84 (1989).
7. Streeter, J. G. *Pl. Physiol.* **90**, 779–782 (1989).
8. Vincent, J. M. & Humphrey, B. A. *Nature* **199**, 149–151 (1963).
9. Nekolla, S., Anderson, C. & Benz, R. *Biophys. J.* **66**, 1388–1397 (1994).
10. McGeoch, M. W. & McGeoch, J. E. M. *Biophys. J.* **66**, 161–168 (1994).
11. Sauvé, R. & Szabo, G. J. *theor. Biol.* **113**, 501–516 (1985).
12. Kolb, H.-A. & Läuger, P. J. *Membr. Biol.* **41**, 167–187 (1978).

baskets for use as substrates in dissociation experiments. Loss of at least 95% of the auxilin (not shown) reduced the rate and extent of uncoating by about half, implying that auxilin, or something associated with it, participates with hsp70c in the dissociation reaction (Fig. 1a). Addition of recombinant histidine-tagged auxilin (H6-auxilin)<sup>11</sup>, together with hsp70c and ATP, restored the original rate and extent of dissociation of baskets (Fig. 1a), thus providing evidence that auxilin is itself the cofactor. The residual cofactor activity of auxilin-depleted preparations is consistent with our previous observation that the cofactor functions even in sub-stoichiometric concentrations<sup>3</sup>. We cannot, however, rule out minor contamination with active auxilin fragments, lacking the epitope for antibody 100/4 or with other cofactors, not recognized by the antibody. To confirm the function of auxilin in a fully defined system, baskets were prepared from purified clathrin in the presence of recombinant mouse brain AP180 (ref. 12) which also served to stabilize the lattice at neutral pH. Only in the simultaneous presence of hsp70c, H6-auxilin and ATP was dissociation detected above background level (Fig. 1b).

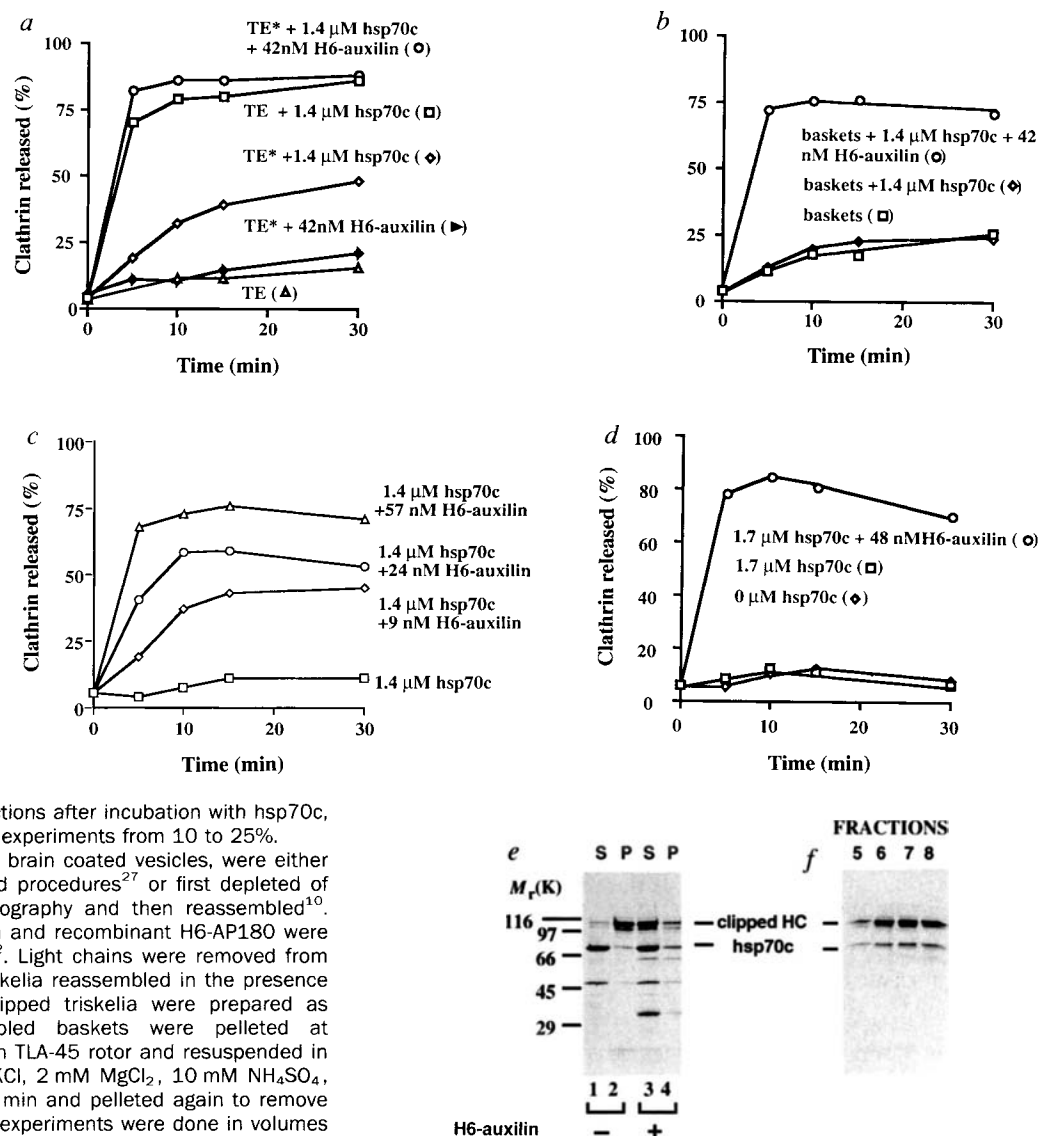
Both clathrin light chains and the N-terminal domain of the heavy chain have previously been implicated in the uncoating reaction<sup>4,5,13</sup>. Because auxilin can bind to baskets devoid of light

chains<sup>14</sup> and the clathrin terminal domain, we re-examined the relation of these elements to the dissociation mechanism. First, the light chains were separated from the heavy chain with thiocyanate<sup>15</sup>. The stripped triskelia were polymerized into baskets in the presence of AP180. Again a sharp rise in the rate and extent of uncoating to a level characteristic of baskets containing light chains occurred only when auxilin and hsp70c were both present (Fig. 1c). Controlled chymotrypsin treatment of baskets selectively degrades light chains<sup>4</sup>. Baskets so treated also required auxilin for their dissociation (not shown). It follows from these results that light chains are not essential for uncoating. Similar observations were made with baskets composed only of trypsinized triskelia, lacking both light chains and the terminal domain of the heavy chain<sup>16</sup>. In the presence of hsp70c but absence of auxilin these baskets proved even more stable than intact, AP180-containing baskets, but when auxilin was included in the reaction they also dissociated rapidly (Fig. 1d, e). Gel filtration revealed that the released triskelia contained 1.5 molecules hsp70c per triskelion (Fig. 1f), but no auxilin (not shown). The composition of these complexes is in agreement with electron micrographs of the clathrin-hsp70c complex, showing triskelia with additional mass near their hubs<sup>17</sup>. Thus we conclude that only the proximal domain of the heavy chain,

FIG. 1 Time course of hsp70c-dependent uncoating of clathrin baskets and its derivatives in the presence and absence of auxilin.

a, Disassembly of baskets (0.3  $\mu$ M triskelia) reconstituted from total coat protein extracts (TE) or from auxilin-depleted extracts (TE\*). Where indicated hsp70c and recombinant H6-auxilin were included<sup>11</sup>. b, Disassembly of AP180-containing clathrin baskets (0.5  $\mu$ M triskelia). c, Disassembly of AP180-containing baskets (0.5  $\mu$ M triskelia) in the absence of light chains; d, hsp70c- and auxilin-dependent uncoating of baskets (0.58  $\mu$ M clipped triskelia) lacking intact clathrin light chains and the N-terminal domain of the heavy chain. e, SDS-PAGE of supernatant (S) and corresponding pellet fractions (P) of trypsin-clipped baskets after 10-min incubation with hsp70c and where indicated with H6-auxilin. f, SDS-PAGE of released clipped triskelia complexed with hsp70c after gel filtration on Superose 6 (Pharmacia). Free hsp70c eluted in fractions 16–19 (not shown). The amount of aggregated clathrin remaining in the pellet fractions after incubation with hsp70c, auxilin and ATP varied for different experiments from 10 to 25%.

METHODS. Coat proteins from cow brain coated vesicles, were either reassembled according to standard procedures<sup>27</sup> or first depleted of auxilin by immunoaffinity chromatography and then reassembled<sup>10</sup>. Baskets composed of pure clathrin and recombinant H6-AP180 were prepared as described previously<sup>12</sup>. Light chains were removed from triskelia with NaSCN<sup>15</sup> and the triskelia reassembled in the presence of recombinant AP180. Trypsin-clipped triskelia were prepared as described elsewhere<sup>5</sup>. Reassembled baskets were pelleted at 109,000g for 20 min in a Beckman TLA-45 rotor and resuspended in buffer C (20 mM HEPES, 25 mM KCl, 2 mM MgCl<sub>2</sub>, 10 mM NH<sub>4</sub>SO<sub>4</sub>, pH 7.0), incubated at 25 °C for 15 min and pelleted again to remove labile coat structures. Dissociation experiments were done in volumes of 75–100  $\mu$ l at 25 °C in buffer C supplemented with 2 mM ATP, 5 mM creatine-phosphate and 5 U ml<sup>-1</sup> phospho-creatine-kinase<sup>18</sup>. Reactions were stopped by adding 1 unit hexokinase and glucose to 5 mM. Dissociated clathrin was separated from baskets by ultracentrifugation and quantified by densitometry after SDS-PAGE using a Molecular Dynamics instrument<sup>10</sup>.



corresponding to the hub of the triskelia, are required for enzymatic dissociation of baskets and that the reaction requires only auxilin, hsp70c and ATP.

We next investigated whether auxilin was required for the binding of hsp70 to clathrin baskets or for the actual uncoating process itself. In this regard we noted that at pH 6.4, where disassembly is inhibited<sup>18</sup>, auxilin aids ATP-dependent association of hsp70c with baskets (Fig. 2a). With only 0.25 bound

auxilins per heavy chain, binding of hsp70c approached saturation at 0.75 bound hsp70c molecules per heavy chain (Fig. 2b). This suggests that either auxilin catalytically induces hsp70c to bind directly to sites on the heavy chain or several hsp70c molecules can attach to a single clathrin-associated auxilin, consistent with our observation that auxilin can induce polymerization of hsp70c<sup>19</sup>. The hsp70c-auxilin-clathrin coat complex did not form in the absence of ATP and requires a supply of ATP to

FIG. 2 Formation of a complex composed of clathrin, auxilin and hsp70c at pH 6.4. *a*, Ap180-containing clathrin baskets (0.44  $\mu$ M) were incubated for 15 min at 25  $^{\circ}$ C in 0.1 M MES buffer, 0.5 mM  $MgCl_2$ , 1 mM EGTA, 0.02%  $Na_2S_2O_8$ , pH 6.4 with hsp70c (1.4  $\mu$ M) and where indicated with 23 nM auxilin in the presence of 2 mM nucleotide of the type indicated below the gel tracks. The baskets were pelleted by ultracentrifugation and aliquots of supernatant and corresponding pellet fractions were analysed by SDS-PAGE. In the absence of clathrin baskets only little hsp70c is found in the pellet (tracks 7 and 8). *b*, Auxilin-dependent binding of hsp70c to AP180-containing clathrin baskets. The molar ratio of bound hsp70c to clathrin is plotted as a function of bound auxilin. METHODS. Baskets (0.37  $\mu$ M triskelia) in the presence of 1.4  $\mu$ M hsp70c were titrated with increasing amounts of H6-auxilin (0–0.14  $\mu$ M). To determine the amount of basket-associated hsp70c and auxilin, the baskets were pelleted by ultracentrifugation and their composition was analysed by SDS-PAGE

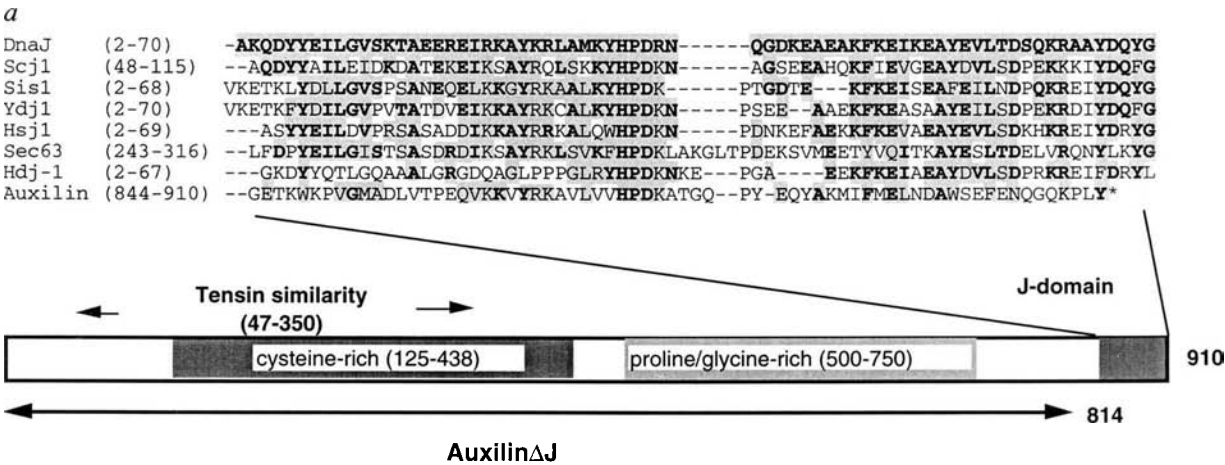
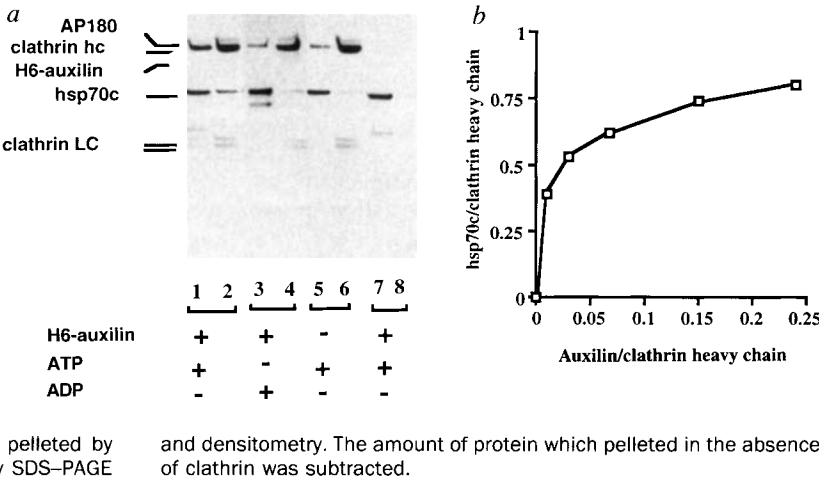
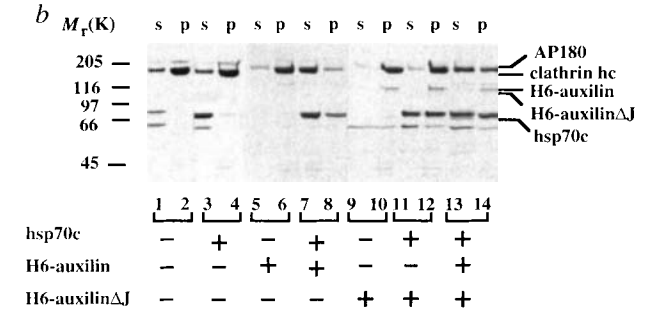


FIG. 3 Auxilin contains a C-terminal J-domain which is required for basket dissociation. *a*, Sequence alignment of J-domains from DnaJ and known eukaryotic DnaJ-like proteins<sup>28</sup> with the C-terminal segment of auxilin. Residues identical (bold type) or homologous to those in DnaJ are shaded. The domain organization of auxilin is shown in the diagram underneath. Numbers denote the positions of sequence- or domain boundaries. AuxilinΔJ refers to a truncated auxilin which terminates at proline 814 and therefore lacks the J-domain. *b*, Properties of truncated auxilin. Ap180-containing baskets were incubated with intact recombinant auxilin or auxilinΔJ and hsp70c in combinations detailed underneath the gel tracks. In the presence of intact recombinant auxilin and hsp70c about 83% of the baskets dissociated (lanes 7 and 8). In contrast recombinant auxilinΔJ at twice the molar concentration did not cause significant release of clathrin (compare lanes 7 and 8 with lanes 11 and 12). When intact auxilin (0.1  $\mu$ M) and auxilinΔJ (0.2  $\mu$ M) were added together to the baskets about 60% of them dissociated (lanes 13 and 14). This is less than expected for the amount of intact auxilin added (compare lanes 7 and 8 with lanes 13 and 14). METHODS. The auxilinΔJ DNA was cut with *Bam*H1 from a pQE-30 vector that contained the full-length auxilin cDNA<sup>11</sup> and inserted into the *Bam*H1 site of an empty pQE-30 vector (Qiagen). For expression of auxilinΔJ we used *E. coli* M15. The bacteria were grown for 3 h at 37  $^{\circ}$ C. Expression was induced with 1 mM isopropyl- $\beta$ -D-thiogalactoside (IPTG)



for 3 h at 25  $^{\circ}$ C. The recombinant protein was purified exactly as described previously for the purification of H6-auxilin<sup>11</sup>. The basket dissociation experiments were done as described in the legend to Fig. 1. The protein concentrations were: 0.5  $\mu$ M clathrin triskelia assembled with AP180 into baskets, 0.1  $\mu$ M intact recombinant auxilin, 0.2  $\mu$ M recombinant truncated auxilin (auxilinΔJ) and 1.5  $\mu$ M hsp70c. The incubation time was 15 min and the temperature 25  $^{\circ}$ C.



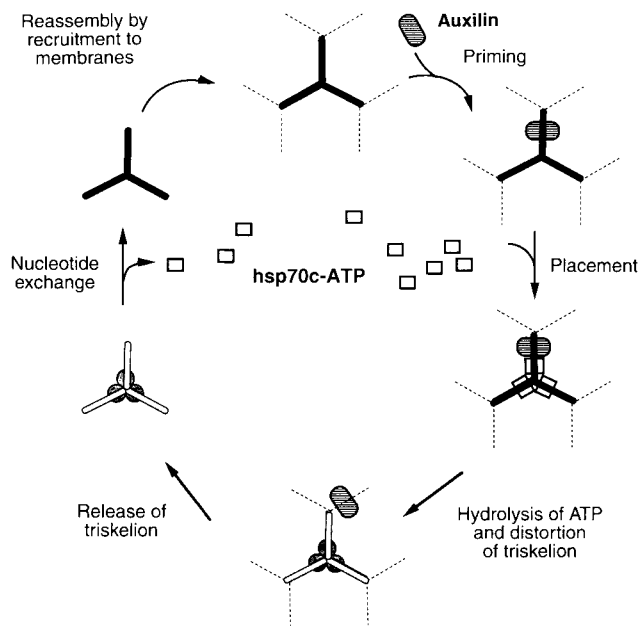


FIG. 4 Model illustrating the function of hsp70c and auxilin in the dissociation of clathrin lattices. Auxilin first attaches to a clathrin molecule in the basket priming it for attachment of hsp70c. Following attachment of hsp70c to the auxilin and clathrin, the auxilin is released and can then attach to another clathrin molecule and initiate another cycle of hsp70c recruitment. Meanwhile the clathrin-bound hsp70c may be prone to self-association<sup>29,30</sup> so that trimerization of the hsp70c occurs on the clathrin causing three monomers of hsp70c to be bound to one triskelion. This trimerization, possibly in association with ATP hydrolysis, may then disrupt clathrin-clathrin interactions causing the release of the triskelion. A similar suggestion has been made previously by Heuser and Steer<sup>30</sup> who referred to this as 'protein polymer competition'. At low pH (6.4), the cohesive forces between the triskelion legs might prevail. Auxilin would still promote the binding of hsp70c to clathrin but, unable to dissociate the triskelion, the hsp70c molecules would eventually detach and engage in a new futile cycle of attachment, ATP hydrolysis and dissociation. Upon nucleotide exchange, hsp70c, containing a bound ATP, dissociates from the triskelion, which can enter a new cycle of coat assembly and disassembly. Hsp70c with bound ATP is symbolized by squares and hsp70c with bound ADP by circles.

maintain itself (not shown), because the rate of ATP hydrolysed, which is uncoupled from basket dissociation, is high at pH 6.4<sup>18</sup>.

The role of auxilin in coat dissociation, specifically that it induces binding of hsp70c to clathrin baskets, bears a striking resemblance to the function of DnaJ proteins: these cooperate with members of the hsp70 family in such diverse functions as protein folding, transfer of proteins across membranes and dissociation of protein complexes<sup>20</sup>. Analysis of the auxilin sequence revealed limited, but significant, homology between the C-terminal residues 846 to 910 and the J-domains of DnaJ-like proteins (Fig. 3a). Overall the similarity between the C-terminal segment of auxilin and known J-domains ranges from 49% (*Escherichia coli* DnaJ) to 62% (yeast Sis1). Residues that are already known to be required for the activity of DnaJ proteins, such as the activation of the ATPase of DnaK<sup>21</sup>, are also strictly conserved in auxilin's J domain. This includes the tripeptide His 33-Pro 34-Asp 35 which occurs in all known J domains. Substitution of His 33 with Gln results in a non-functional DnaJ<sup>22</sup>. Similarly the yeast DnaJ homologue Sec63p which cooperates with the luminal ER hsp70 homologue Kar2p (BiP) in protein transport does not tolerate substitutions in positions that correspond to Pro 34 and Asp 35, respectively<sup>23</sup>.

In auxilin the tripeptide starts in position 874. A recent nuclear magnetic resonance (NMR) study showed that the J domain of *E. coli* is built from three helices which form a hydrophobic core<sup>24</sup>. The hydrophobic residues involved in core formation are

conserved in J domains. With the exception of Lys 847 this is also true for the hydrophobic residues in auxilin, although some of the residues are substituted by homologous ones. In most, but not all DnaJ-like proteins the J-domain is located at the N terminus<sup>20,25</sup>. It is usually separated by a G/F-rich segment from a C-rich domain and other less conserved regions. In auxilin the C-terminal J-domain is preceded by a G/P-rich domain. Auxilin also contains a C-rich domain, which does not, however, conform to the CXXGXGXG repeat found in a number of DnaJ-like proteins<sup>20,25</sup>. There is already strong evidence that the J-domain of a given DnaJ homologue serves not only as a general hsp70 interaction domain but that it also determines the hsp70 partner<sup>26</sup>. Thus, some of the observed sequence differences between J-domains are likely to encode the information for associating with a specific hsp70.

To prove directly an involvement of auxilin's J-domain in basket dissociation we constructed a truncated form of auxilin from which the entire J-domain had been deleted and expressed it in *E. coli*. The truncated auxilin (auxilinΔJ) which terminated at Pro 814, still interacted with clathrin, but it failed, as expected, to support hsp70c in basket dissociation (Fig. 3b).

The involvement of a DnaJ-like protein in the uncoating process not only changes our previous perception of this reaction (Fig. 4) but it also provides us with a simple biochemical system to investigate the cooperation of hsp70c with a specific partner protein that defines the molecular target of the chaperone. □

Received 6 June; accepted 4 October 1995.

1. Brodsky, F. M. *Science* **242**, 1396–1402 (1988).
2. Goldstein, J. L., Brown, M. S., Anderson, R. G., Russell, D. W. & Schneider, W. J. A. *Rev. Cell Biol.* **1**, 1–39 (1985).
3. Prasad, K., Barouch, W., Greene, L. & Eisenberg, E. *J. biol. Chem.* **268**, 23758–23761 (1993).
4. Schmid, S. L., Braell, W. A., Schlossman, D. M. & Rothman, J. E. *Nature* **311**, 228–231 (1984).
5. Schmid, S. L. & Rothman, J. E. *J. biol. Chem.* **260**, 10050–10056 (1985).
6. Hartl, F. U. *Semin. Immun.* **3**, 5–16 (1991).
7. Georgopoulos, C. & Welch, W. J. A. *Rev. Cell Biol.* **9**, 601–634 (1993).
8. Gething, M. J. & Sambrook, J. *Nature* **355**, 33–45 (1992).
9. Ahle, S. & Ungewickell, E. *J. Cell Biol.* **111**, 19–29 (1990).
10. Lindner, R. & Ungewickell, E. *J. biol. Chem.* **267**, 16567–16573 (1992).
11. Schröder, S. et al. *Eur. J. Biochem.* **228**, 297–304 (1995).
12. Norris, F. A., Ungewickell, E. & Majerus, P. W. *J. biol. Chem.* **270**, 214–217 (1995).
13. DeLuca-Flaherty, C., McKay, D. B., Parham, P. & Hill, B. L. *Cell* **62**, 875–887 (1990).
14. Lindner, R. & Ungewickell, E. *Biochemistry* **30**, 9097–9101 (1991).
15. Winkler, F. K. & Stanley, K. K. *EMBO J.* **2**, 1393–1400 (1983).
16. Schmid, S. L., Matsumoto, A. K. & Rothman, J. E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 91–95 (1982).
17. Prasad, K., Heuser, J., Eisenberg, E. & Greene, L. *J. biol. Chem.* **269**, 6931–6939 (1994).

18. Braell, W. A., Schlossman, D. M., Schmid, S. L. & Rothman, J. E. *J. Cell Biol.* **99**, 734–741 (1984).
19. King, C., Eisenberg, E. & Greene, L. *J. biol. Chem.* **270**, 22535–22540 (1995).
20. Cyr, D. M., Langer, T. & Douglas, M. G. *Trends biochem. Sci.* **19**, 176–181 (1994).
21. Liberek, K., Marszałek, J., Ang, D., Georgopoulos, C. & Zyllicz, M. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2874–2878 (1991).
22. Wall, D., Zyllicz, M. & Georgopoulos, C. *J. biol. Chem.* **269**, 5446–5451 (1994).
23. Feldheim, D., Rothblatt, J. & Schekman, R. *Molec. cell. Biol.* **12**, 3288–3296 (1992).
24. Szyperki, T., Pellicchia, M., Wall, D., Georgopoulos, C. & Wuthrich, K. *Proc. natn. Acad. Sci. U.S.A.* **91**, 11343–11347 (1994).
25. Silver, P. A. & Way, J. C. *Cell* **74**, 5–6 (1993).
26. Schlenstedt, G., Harris, S., Risse, B., Lill, R. & Silver, P. A. *J. Cell Biol.* **129**, 979–988 (1995).
27. Keen, J. H., Willingham, M. C. & Pastan, I. *Cell* **16**, 303–312 (1979).
28. Caplan, A. J., Cyr, D. M. & Douglas, M. G. *Molec. Biol. Cell* **4**, 555–563 (1993).
29. Schmid, S. L., Braell, W. A. & Rothman, J. E. *J. biol. Chem.* **260**, 10057–10062 (1985).
30. Heuser, J. & Steer, C. J. *J. Cell Biol.* **109**, 1457–1466 (1989).

ACKNOWLEDGEMENTS. W. Gratzer and L. Traub are thanked for their helpful comments. M. Feldmeier, S. Vora, J. Kilheeneey and A. Ungewickell are thanked for able technical assistance. This work was supported generously by start-up funds from the Department of Pathology at Washington University, St Louis, MO.

## Actin-based motility of vaccinia virus

Sally Cudmore\*, Pascale Cossart†, Gareth Griffiths\* & Michael Way\*‡

\* Cell Biology Programme, European Molecular Biology Laboratory, 1 Meyerhofstrasse, Heidelberg 69117, Germany

† Unité des Interactions Bactéries Cellules, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris Cedex 15, France

‡ To whom correspondence should be addressed

**THE role of the cytoskeleton during viral infection is poorly understood. Here we show, using a combination of mutant and drug studies, that the intracellular enveloped form of vaccinia virus is capable of inducing the formation of actin tails that are strikingly similar to those seen in *Listeria*, *Shigella* and *Rickettsia* infections. Analysis using video microscopy reveals that single viral particles are propelled *in vivo* on the tip of actin tails, at a speed of  $2.8 \mu\text{m min}^{-1}$ . On contact with the cell surface, virus particles extend outwards on actin projections at a similar rate, to contact and infect neighbouring cells. Given the similarities between the motility of vaccinia virus and bacterial pathogens, we suggest that intracellular pathogens have developed a common mechanism to exploit the actin cytoskeleton as a means to facilitate their direct spread between cells.**

Vaccinia virus, a close relative of the smallpox virus, has a genome of ~190 kilobase pairs (kbp) encoding over 260 potential open reading frames (ORFs), and so is one of the largest and most complex viruses known<sup>1,2</sup>. It is different from other viral families in that it has two distinct infectious forms, the intracellular mature virus (IMV) and the extracellular enveloped virus (EEV). IMV begin to appear about 4 hours post-infection, and are only released when the cell lyses, presumably as a result of the cytotoxic effects of infection. Alternatively, a small proportion of IMV can become wrapped by membrane cisternae derived from the *trans* Golgi network to form the intracellular

enveloped virus (IEV), which is first seen about 6 hours post-infection<sup>3,4</sup>. Infectious EEV are thought to be released when the outermost membrane of the IEV fuses with the plasma membrane<sup>3,5</sup>.

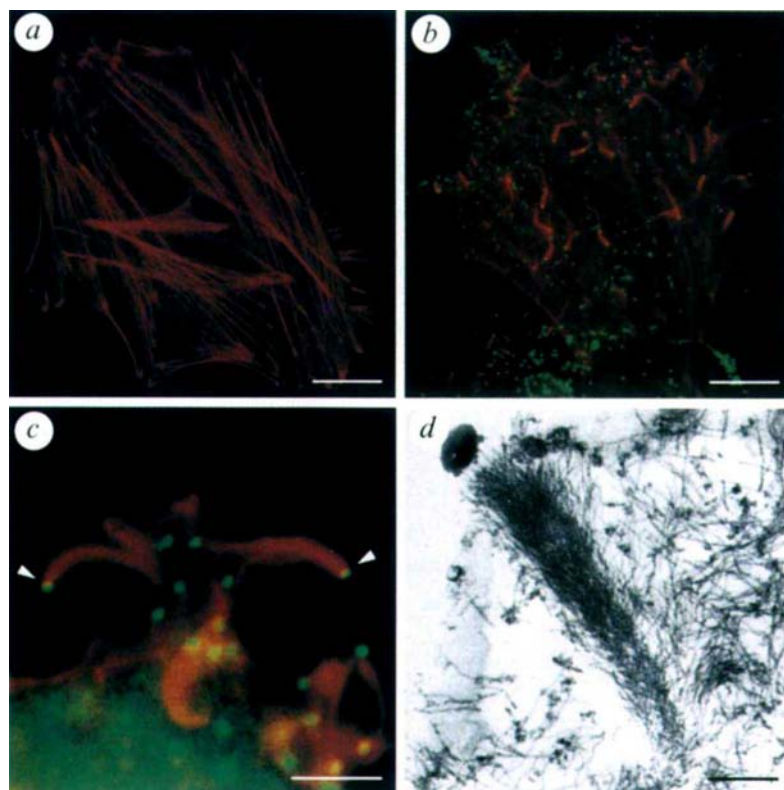
High-voltage electron microscopy has previously shown that, late in infection, vaccinia induces the formation of large microvilli with single virus particles at their tip<sup>6</sup>. Subsequent studies confirmed that actin was present in these microvilli<sup>7,10</sup>. We found in all cell types tested that ~6 hours post-infection the vaccinia virus strain WR begins to induce the formation of actin-containing structures that are very reminiscent of the actin tails formed during bacterial infection<sup>11–15</sup> (Fig. 1). Infection also results in a significant decrease in the number of actin stress fibres in the cell. Actin tail formation occurs from single viral particles, which always remain at the tip of these structures (Fig. 1b). Actin tails are initially seen in the interior of the cell, but as infection proceeds increasing numbers are seen projecting from the cell surface (Fig. 1c). Intracellular actin tails, which taper away from the virus, are typically 6.4–9.6  $\mu\text{m}$  long ( $n=64$ ), whereas projections can reach lengths of over 20  $\mu\text{m}$ . Electron microscopy confirms that single viral particles are in direct contact with actin filaments (Fig. 1d), which is consistent with the overlap between actin and virus particles seen in fluorescence images (Fig. 1c).

Although actin is necessary in the later stages of infection, it is not essential during the initial stages of infection or IMV assembly, as these processes occur in cells infected in the presence of cytochalasin D (data not shown). However, the presence of cytochalasin D prevents the formation of tails and subsequent release of IEV, although actin tails are able to form rapidly when cytochalasin D is removed (data not shown).

The time of appearance and number of viral particles inducing actin tails (5–15%) is consistent with the formation of IEV during vaccinia virus infection of HeLa cells<sup>3</sup>. We tested whether IEV, and not IMV, induces the formation of actin tails using both the vaccinia mutant, vRB 12, which is highly deficient in the production of IEV<sup>16</sup> and the drug  $\text{N}_1$ -isonicotinoyl- $\text{N}_2$ -3-methyl-4-chlorobenzoylhydrazine (IMCBH), which blocks IEV formation<sup>16–20</sup>. In both cases, IMV is confined to the perinuclear

**FIG. 1** Actin tails in HeLa cells infected with vaccinia virus. **a**, Uninfected control HeLa cell which shows a typical actin stress fibre staining. Scale bar, 50  $\mu\text{m}$ . **b**, Infected cells showing intracellular actin tails (red) and viral particles (green). Scale bar, 50  $\mu\text{m}$ . **c**, Viral particles projecting from the cell surface on actin tails show a clear overlap between the viral particle and actin (white arrowheads). Scale bar, 10  $\mu\text{m}$ . **d**, Electron micrograph of an actin tail decorated with S1 myosin showing a direct contact between actin filaments and the virus particle. Scale bar, 400 nm.

**METHODS.** **a–c**, HeLa cells were infected with vaccinia virus strain WR at a multiplicity of infection of one plaque-forming unit per cell. The cells were fixed 8 h after infection and stained as described<sup>29</sup> with rhodamine phalloidin and the anti-vaccinia antibody, C3, which recognizes the peripheral membrane protein of relative molecular mass 14,000 ( $M_r$  14K) of the IMV<sup>30</sup>, followed by FITC goat anti-mouse. **d**, Infected HeLa cells were made permeable with 1 U  $\text{ml}^{-1}$  streptolysin O (SLO) in SLO buffer (150 mM sucrose, 50 mM KAc, 5 mM MgAc, 10 mM HEPES, 1 mM DTT) on ice for 10 min followed by a 15-min incubation at 37 °C. The cells were washed once in 50 mM phosphate buffer (pH 6.8) and incubated with 2 mg  $\text{ml}^{-1}$  S1 myosin for 30 min on ice. After two washes in ice-cold PBS, the cells were fixed in 1% glutaraldehyde and 2% tannic acid for 30 min. The specimens were then post-fixed in 1% osmium on ice for 30 min, dehydrated in ethanol, and embedded in EPON.



region of the cell, and no actin tails are seen while the actin cytoskeleton appears normal (Fig. 2). Subsequent removal of IMCBH results in the appearance of actin tails with associated viral particles (Fig. 2c).

We observed individual viral particles *in vivo* by video microscopy (Fig. 3). Phase-dense structures, the actin tails, were only observed behind moving viral particles and were completely absent from vRB12-infected cells. The number of moving particles and tails is consistent with fluorescence images of fixed cells, although the tails generally appear shorter. Movement within cells appears to be random with an average speed of  $2.8 \mu\text{m min}^{-1}$  (s.d.  $0.5 \mu\text{m min}^{-1}$ ,  $n=44$ ), which is similar to that observed for *Listeria*<sup>21,22</sup>. On reaching the cell surface, virus particles project up to  $20 \mu\text{m}$  at the same rate (Fig. 3). Projections are always much larger than the normal cell-surface microvilli, and presumably correspond to the virally induced projections previously described<sup>6,7</sup>.

Cytochalasin D has previously been shown to prevent the release of IEV, which suggests that actin is involved in this process<sup>23</sup>. Our immunofluorescence images clearly show viral-tipped actin projections extending into uninfected cells (Fig. 4a). Electron microscopy also shows virus particles at the tip of a dense actin network projecting from one cell into another (Fig. 4b). Taken together, our results suggest that IEV are able to recruit and exploit actin to facilitate their spread from cell to cell. Given that the rates at which intracellular tails move and projections extend are identical, it is likely that these two events are driven by the same mechanism. However, the stability of actin in tails and projections is very different: the actin at the distal end of the intracellular tail is unstable and depolymerizes as the tail continues to extend, but it is stable in the elongating projection. This stability may be explained in part by membrane interactions in the projection that are absent in intracellular tails.

FIG. 2 IEV induces actin tail formation. **a**, HeLa cells infected with a mutant vaccinia strain, vRB 12. **b**, Cells infected with wild-type WR in the presence of IMCBH. The virus is concentrated in the perinuclear area of both cells and there are no actin tails. **c**, Viral particles are seen on actin tails 1 h after IMCBH removal. Scale bar,  $50 \mu\text{m}$ . **METHODS**. Cells were infected with WR or vRB12 as described above. IMCBH was added to a final concentration of  $10 \mu\text{g ml}^{-1}$  1 h post-infection. At 8 h post-infection cells were washed 5 times with PBS and fixed 1 h after washout.

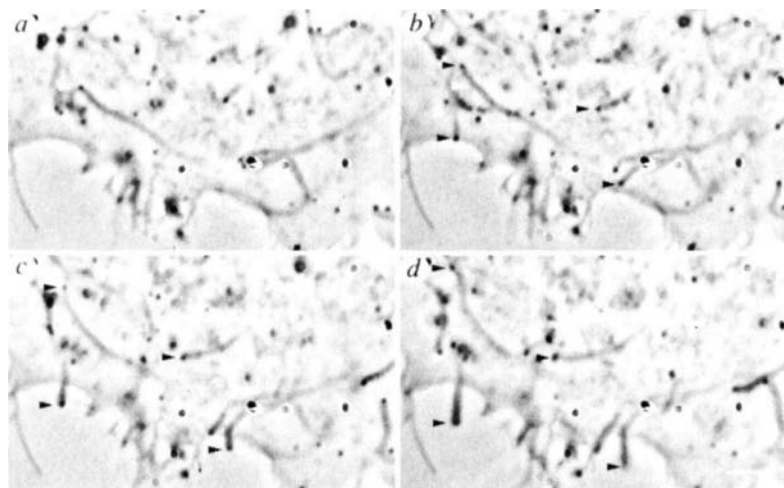
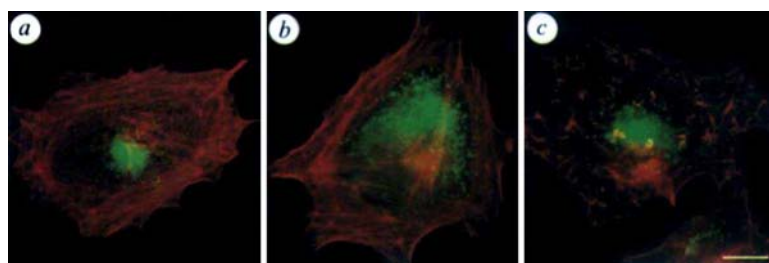
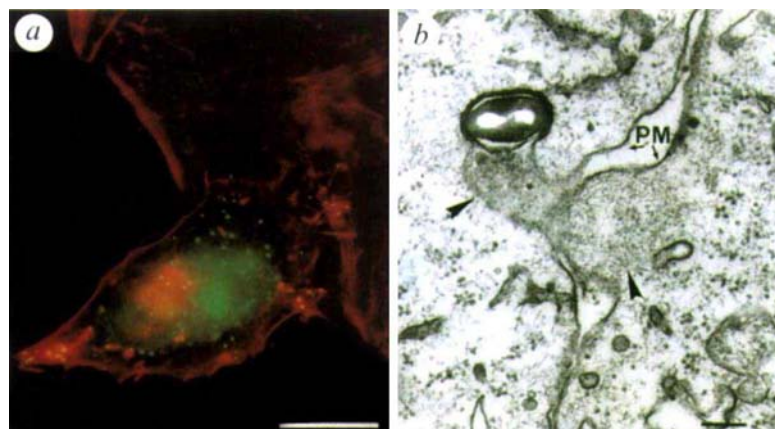


FIG. 3 Virus motility *in vivo*. Frame stills 40 s apart, from the video microscope, showing virus particles moving through the cell at the tip of actin tails and inducing the formation of protrusions from the cell surface. White arrowheads are fixed reference points, and black arrows indicate moving virus particles. Scale bar,  $3.0 \mu\text{m}$ . **METHODS**. Cells were infected with WR and viewed 10 h post-infection with a Zeiss axiovert microscope using a Sony CCD IRIS camera.

FIG. 4 Spread of vaccinia virus from cell to cell. **a**, Virus particles moving from an infected cell to an uninfected cell at the tip of actin-rich projections. Scale bar,  $50 \mu\text{m}$ . **b**, Electron micrograph of a virus particle moving into a neighbouring cell. Black arrowheads show actin, and PM indicates the plasma membrane of the two cells. Scale bar,  $400 \text{ nm}$ .

**METHODS**. **a**, Cells were infected in the presence of IMCBH. Then, 5 h post-infection, the cells were washed once with warm PBS + IMCBH and uninfected cells were seeded on the same coverslips; 3 h later the IMCBH was washed out, and the cells were fixed after 1 h and stained as described above. **b**, Infected cells were fixed in 1% glutaraldehyde,  $0.5 \text{ mg ml}^{-1}$  saponin, 2% tannic acid, 50 mM KCl, 5 mM  $\text{MgCl}_2$  for 30 min, post-fixed in 1% osmium on ice for 30 min, dehydrated in ethanol, and embedded in EPON.





The intimate relationship between the IEV and the actin cytoskeleton suggests that the virus contains protein(s) that can associate directly or indirectly with actin. In *Listeria* species and *Shigella*, the proteins ActA, IactA and IcsA are critical for the nucleation of actin filaments at the bacterium surface<sup>15,24,26</sup>. Sequence comparisons of the vaccinia virus genome with ActA, IactA or IcsA failed to identify any ORF that contains significant homologous sequences. One possible candidate for viral interactions with actin is the A42R ORF, which is 32.1% identical to human profilin<sup>27</sup>. However, deletion of A42R has previously shown that it is not required for infectivity or release

of mature virions<sup>27</sup>. Moreover, we found that vaccinia virus lacking A42R is still able to form actin tails and projections (data not shown), which is consistent with the recent observation that profilin is not required for actin-based motility of *Listeria* in the cell-free *Xenopus* system<sup>28</sup>. Whether profilin is involved in vaccinia motility will require further characterization, as the virus may be able to utilize cellular profilin isoforms. Given the similarities between the motility of bacterial pathogens and vaccinia virus, we suggest that intracellular pathogens have developed a common mechanism to exploit the actin cytoskeleton for their own purposes. □

Received 25 August; accepted 6 October 1995.

- Goebel, S. J. *et al.* *Virology* **179**, 247–266 (1990).
- Moss, B. in *Virology* (eds Fields, B. N. *et al.*) 2079–2111 (Raven, New York, 1990).
- Payne, L. G. *J. gen. Virol.* **50**, 89–100 (1980).
- Schmelz, M. *et al.* *J. Virol.* **68**, 130–147 (1994).
- Morgan, C. *Virology* **73**, 43–58 (1976).
- Stokes, G. V. *J. Virol.* **18**, 636–643 (1976).
- Hiller, G., Weber, K., Schneider, L., Parajsz, C. & Jungwirth, C. *Virology* **98**, 142–153 (1979).
- Hiller, G. & Weber, K. *J. Virol.* **44**, 647–657 (1982).
- Hiller, G., Jungwirth, C. & Weber, K. *Exp. Cell Res.* **132**, 81–87 (1981).
- Krempien, U. *et al.* *Virology* **113**, 556–564 (1981).
- Pollard, T. D. *Curr. Biol.* **5**, 837–840 (1995).
- Cossart, P. *Curr. Opin. Cell Biol.* **7**, 94–101 (1995).
- Tilney, L. G. & Portnoy, D. A. *J. Cell Biol.* **109**, 1597–1608 (1989).
- Heinzen, R. A., Hayes, S. F., Peacock, M. G. & Hackstead, T. *Infect. Immun.* **61**, 1926–1935 (1993).
- Bernardini, M. L., Mounier, J., d'Hauteville, H., Coquis-Rondon, M. & Sansonetti, P. J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 3867–3871 (1989).
- Blasco, R. & Moss, B. *J. Virol.* **66**, 4170–4179 (1992).
- Hiller, G., Eibl, H. & Weber, K. *J. Virol.* **39**, 903–913 (1981).
- Kato, N., Eggers, H. J. & Rolly, H. *J. exp. Med.* **129**, 795–808 (1969).

- Payne, L. G. & Kristenson, K. *J. Virol.* **32**, 614–622 (1979).
- Schmutz, C., Payne, L. G., Gubser, J. & Witte, R. *J. Virol.* **65**, 3435–3442 (1991).
- Dabiri, G. A., Sanger, J. M., Portnoy, D. A. & Southwick, F. S. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6068–6072 (1990).
- Theriot, J. A., Mitchison, T. J., Tilney, L. G. & Portnoy, D. A. *Nature* **357**, 257–260 (1992).
- Payne, L. G. & Kristenson, K. *Archs Virol.* **74**, 11–20 (1982).
- Kocks, C. *et al.* *Cell* **68**, 521–531 (1992).
- Gouin, E., Dehoux, P., Mengaud, J., Kocks, C. & Cossart, P. *Infect. Immun.* **63**, 2729–2737 (1995).
- Domann, E. *et al.* *EMBO J.* **11**, 1981–1990 (1992).
- Blasco, R., Cole, N. B. & Moss, B. *J. Virol.* **65**, 4598–4608 (1991).
- Marchand, J.-B. *et al.* *J. Cell Biol.* **130**, 1–13 (1995).
- Herzog, M., Draeger, A., Ehler, E. & Smali, J. V. *Cell Biology: A Laboratory Handbook* (Academic, San Diego, 1994).
- Rodriguez, J. F., Janeczko, R. & Esteban, M. *J. Virol.* **56**, 482–488 (1985).

ACKNOWLEDGEMENTS. We thank Y. Rivière, V. David and E. Gouin for help in preliminary experiments; R. Blasco for vRB 12 and profilin null virus; D. Klaus and M. Esteban for the S1 myosin and C3 antibody, respectively; C. Dotti for help and suggestions concerning video microscopy; I. Reckmann for vaccinia virus preparations; S. Schleich for help with electron microscopy; and A. Hyman, K. Simons, M. Zerial, R. Parton and E. Karsenti for suggestions and comments on the manuscript.

## Microtubule nucleation by $\gamma$ -tubulin-containing rings in the centrosome

Michelle Moritz, Michael B. Braunfeld\*, John W. Sedat\*, Bruce Alberts† & David A. Agard\*

Department of Biochemistry & Biophysics and  
\* Howard Hughes Medical Institute, University of California at  
San Francisco, 513 Parnassus Avenue, San Francisco,  
California 94143-0448, USA

THE microtubule cytoskeleton of animal cells does not assemble spontaneously, but instead requires the centrosome. This organelle consists of a pair of centrioles surrounded by a complex collection of proteins known as the pericentriolar material (PCM)<sup>1</sup>. The PCM is required for microtubule nucleation<sup>2</sup>. The minus, or slow-growing, ends of microtubules are embedded in the PCM and the plus, or fast-growing, ends project outwards into the cytoplasm during interphase, or into the spindle apparatus during mitosis.  $\gamma$ -Tubulin is the only component of the PCM that is so far implicated in microtubule nucleation<sup>3–6</sup>. Here we use immuno-electron microscopic tomography to show that  $\gamma$ -tubulin is localized in ring structures in the PCM of purified centrosomes without microtubules. When these centrosomes are used to nucleate microtubule growth,  $\gamma$ -tubulin is localized at the minus ends of the microtubules. We conclude that microtubule-nucleating sites within the PCM are ring-shaped templates that contain multiple copies of  $\gamma$ -tubulin.

We previously reported that the PCM of isolated *Drosophila* centrosomes that lack microtubules (MTs) contains hundreds of

TABLE 1 Distribution of gold particles in PCM of centrosomes with and without microtubules

|                             | Number of gold particles<br>(% total particles) |                       |            |
|-----------------------------|-------------------------------------------------|-----------------------|------------|
|                             | At (–) ends<br>of MTs                           | Ambiguous<br>position | Total      |
| (a) Centrosomes with MTs    |                                                 |                       |            |
| All gold                    | 77 (65%)                                        | 41 (35%)              | 118 (100%) |
| Single gold particles       | 2 (2%)                                          | 7 (6%)                |            |
| Clusters of 2               | 12 (10%)                                        | 8 (7%)                |            |
| Clusters of 3               | 15 (13%)                                        | 12 (10%)              |            |
| Clusters of 4               | 24 (20%)                                        | 8 (7%)                |            |
| Clusters of >4              | 24 (20%)                                        | 6 (5%)                |            |
| (b) Centrosomes without MTs | At ring<br>structures                           | Ambiguous<br>position | Total      |
| All gold                    | 68 (69%)                                        | 30 (31%)              | 98 (100%)  |
| Single gold particles       | 3 (3%)                                          | 6 (6%)                |            |
| Clusters of 2               | 12 (12%)                                        | 8 (8%)                |            |
| Clusters of 3               | 12 (12%)                                        | 3 (3%)                |            |
| Clusters of 4               | 36 (37%)                                        | 0 (0%)                |            |
| Clusters of >4              | 5 (5%)                                          | 13 (13%)              |            |

The gold particles were randomly chosen in each reconstruction and their association with identifiable structures (MTs or rings) was then determined by stepping through the relevant projections from the reconstruction, which reveals the three-dimensional structure of objects.

ring-like structures that are 25–30 nm in diameter and 10–13 nm thick. In addition, we found that the minus ends of MTs in centrosomes with regrown asters appear to be blunt and arise abruptly, inasmuch as they are not associated with any discernible structure<sup>7</sup>. We suggested then that the ring structures are MT-nucleating sites. Because their diameter is similar to that of MTs, they could be invisible at the minus ends of MTs in centrosomes with regrown asters. If these suggestions are correct,

† Present address: National Academy of Sciences, 2101 Constitution Avenue NW, Washington, DC 20418, USA.



we expect gold-labelled antibodies against  $\gamma$ -tubulin to stain the rings in centrosomes without MTs and to stain the minus ends of MTs in centrosomes with MTs. In this study, we have tested this idea, determining the three-dimensional structure of the gold-labelled centrosomes by immuno-electron microscopic (EM) tomography<sup>7-9</sup>.

For the three-dimensional reconstructions, purified *Drosophila* centrosomes were stained with gold-labelled antibodies, embedded, cut into semi-thick (0.7  $\mu\text{m}$ ) sections and examined in the EM (Fig. 1). Because these sections can contain entire *Drosophila* centrosomes (which vary in diameter from  $\sim 0.2$  to  $1 \mu\text{m}^2$ ), a complete, high-resolution ( $\sim 6$ – $8 \text{ nm}$ ) three-dimensional image of an intact centrosome can be reconstructed by EM tomography. In some cases, individual MTs can then be followed through such reconstructions to their minus ends, allowing the origins of MTs to be identified unambiguously<sup>7</sup>.

We have used this approach to determine the location of  $\gamma$ -tubulin in the centrosome under two different conditions: first, in the centrosome, as isolated, without any MTs, and second, after the same centrosomes have been mixed with pure tubulin and used to grow asters composed of hundreds of MTs. We present the results for centrosomes with asters in Fig. 1: the centrosomes are abundantly and specifically labelled with gold throughout the PCM (Fig. 1a, white dots, arrowed). In enlarged views of selected regions of the PCM, it can be seen that the minus ends of MTs are labelled with clusters of gold, which are often arranged in semicircles or rings which follow the shape of the MT minus end (Fig. 1b). When we analysed 100 randomly chosen gold particles, we found that 65% of them were associated with MT minus ends; the remaining 35% could not be localized unambiguously because of the close and complicated arrangement of MTs within the PCM. In addition, the majority (62%)

of the gold particles that were associated with MT minus ends were in clusters of four or more (Table 1a). In no case was the gold seen along the length of MTs, even in regions where the MTs were embedded in the PCM. As controls, centrosomes were incubated with the gold-labelled secondary antibodies alone, or with a primary antibody (a gift from C. Field) that recognizes the non-centrosomal protein, peanut (a contractile-ring component), followed by these antibodies. Sections of  $0.2 \mu\text{m}$  were prepared from control samples and scanned by conventional transmission EM for gold particles. In both cases, the number of gold particles found to label the control sample was less than 1% of the number found in samples that had been incubated with antibodies recognizing  $\gamma$ -tubulin. In addition, the gold particles on control samples were never found in clusters (data not shown).

The presence of multiple gold particles at the ends of single MTs suggests that there are multiple  $\gamma$ -tubulin molecules at these positions, in support of a model that postulates that a  $\gamma$ -tubulin molecule sits at the minus end of each of the 13 protofilaments in a MT<sup>3</sup>. In addition, the semicircular or circular arrangement of the gold particles closely follows the shape of the MTs. However, our current results do not allow us to ascertain the number of  $\gamma$ -tubulin molecules in the nucleating structure (13 are predicted).

The results for centrosomes without MTs are presented in Fig. 2. In unlabelled centrosomes, numerous ( $\sim 400$  to  $800$ ) ring structures are visible throughout the PCM<sup>7</sup>. When centrosomes without MTs were immuno-gold-labelled to detect  $\gamma$ -tubulin, the gold particles were distributed in clusters throughout the PCM (Fig. 2a, white dots, arrowed). In enlarged views, we found that approximately 60% of the rings that are visible in the reconstruction are associated with at least one gold particle. More-

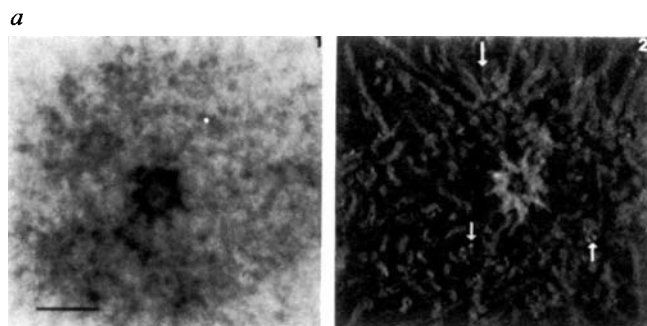
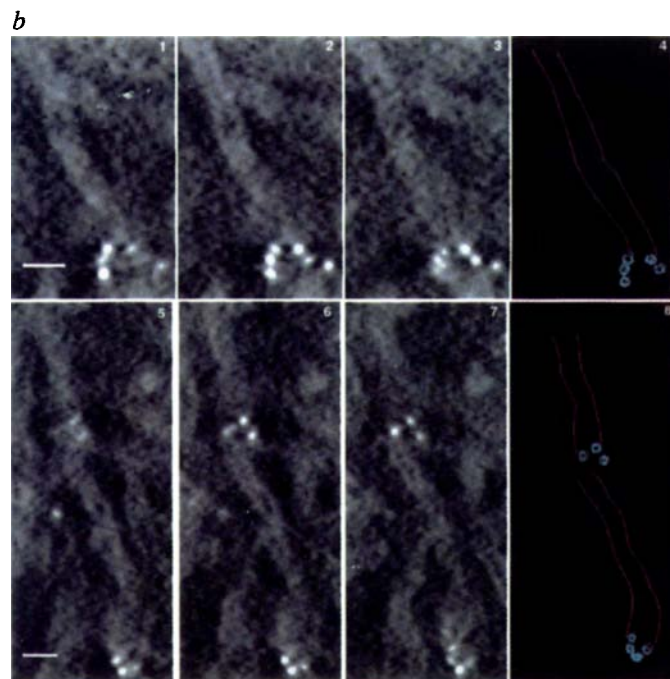


FIG. 1 Tomographic reconstruction of a centrosome with MTs after immuno-gold labelling to detect  $\gamma$ -tubulin. a, Panel 1 shows an example of raw data used to generate the reconstruction shown in panel 2 (Magnification,  $\times 10,400$ ; scale bar,  $200 \text{ nm}$ ); panel 2 shows a selected projection (representing a slice  $2.75 \text{ nm}$  thick) from the computer-generated reconstruction (out of a total of 179). White dots (arrowed) are  $6\text{-nm}$  gold particles. b,  $\gamma$ -Tubulin is localized in semicircles and rings at the minus ends of MTs. Panels 1–3 and 5–7: enlarged views of two examples of gold-labelled MT minus ends taken from the reconstruction in a. Selected projections stepping through the three-dimensional structure are shown ( $2.75 \text{ nm}$  per step). Bar,  $30 \text{ nm}$ . Panels 4 and 8: models of the relevant MTs and gold particles seen in panels 1–3 and 5–7. Red, MT outlines; aqua,  $6\text{-nm}$  gold particles.

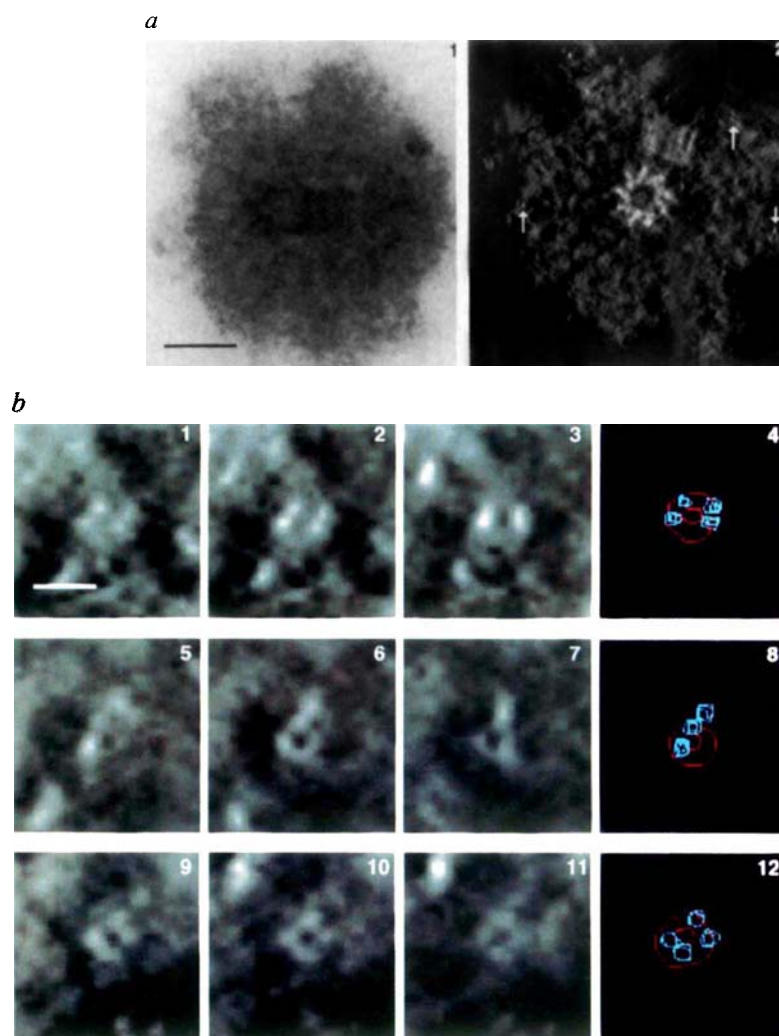
**METHODS.** Centrosomes were purified from 0–3.5-hr *Drosophila* embryos, MTs were regrown on them, and the resulting asters were then fixed and sedimented onto coverslips as described in Ref. 7. For immuno-gold labelling of  $\gamma$ -tubulin, nonspecific binding sites were saturated by a 20-min incubation in Block, consisting of PBS<sup>7</sup> plus 0.5% fish gelatin, 0.8% bovine serum albumin, 0.1% Tween-20, and 0.5 M NaCl. Samples were then incubated for 3 h with an affinity-purified rabbit polyclonal antibody specific for  $\gamma$ -tubulin, diluted 1:500 in Block, which was raised against a peptide corresponding to the unique 17-amino-acid C terminus of the maternal form of *Drosophila*  $\gamma$ -tubulin and affinity-purified on the same peptide (Y. Zheng, personal communication). Samples were washed quickly in Block three times and then



three times for 5 min in PBS plus 0.1% Tween-20. They were then incubated for 3 h, with gentle rocking, in  $6\text{-nm}$ -gold-labelled, affinity-purified goat-anti-rabbit-IgG (EM grade, Jackson Immunological Laboratories) diluted 1:25 in Block. Washes were repeated, and samples were then washed into PBS and fixed in 1% glutaraldehyde in PBS for 15 min. Further preparation of samples for EM was as described in ref. 7. The tomographic data set consisted of 99 views covering a tilt range of  $-63.75^\circ$  to  $+58.75^\circ$  at  $1.25^\circ$  intervals. Images were recorded as  $480 \times 480$  pixel arrays ( $2 \times 2$  binning) on a Gatan photometric-UCSF CCD, giving a final magnification of  $2.75 \text{ nm}$  per pixel. Data were collected automatically using our EMAC data collection system and processed using the EMCAT suite of software<sup>13</sup> as in ref. 7.

FIG. 2 Tomographic reconstructions of purified centrosomes without MTs after immuno-gold labelling to detect  $\gamma$ -tubulin. The  $\gamma$ -tubulin is localized to ring structures scattered throughout the PCM. **a**, Panel 1: An example of the raw data used for generating the reconstruction shown in panel 2 (magnification,  $\times 10,400$ ; bar, 200 nm); panel 2: a selected projection (representing a slice 2.75 nm thick) from the computer-generated reconstruction (out of a total of 143). White dots (arrowed) are 6-nm gold particles. **b**, The  $\gamma$ -tubulin is found throughout the PCM in ring structures 25–30 nm in diameter and 10–13 nm thick. Panels 1–3, 5–7 and 9–11: enlarged views of three examples of gold-labelled ring structures taken from the reconstruction in **a**. Selected projections stepping through the three-dimensional structure are shown (2.75 nm per step). Panels 4, 8, 12: models showing the relevant rings and gold particles in panels 1–3, 5–7 and 9–11. Red, ring structure; aqua, 6-nm gold particles. Bar, 30 nm.

**METHODS.** Samples were prepared as described for Fig. 1 and in ref. 7, except that MTs were not regrown on the purified centrosomes. The tomographic data set consisted of 105 views covering a tilt range of  $-65^\circ$  to  $+65^\circ$  at  $1.25^\circ$  intervals.



over, 69% of the gold particles were clearly associated with the ring structures, and most (60%) of this gold is in clusters of 4 or more (Fig. 2b and Table 1b). The clusters of gold associated with the rings are often arranged in semicircles and circles that correspond to the shape of the rings. We conclude that MT-nucleating sites consist of ring-shaped structures containing multiple copies of  $\gamma$ -tubulin. The general idea that some sort of ring could serve as the template for MT growth has been around for some time<sup>10,11</sup>.

The conclusion that the MT-nucleating site in the centrosome is a ring complex is strongly substantiated by work described in an accompanying paper<sup>12</sup>, in which a ring-shaped complex of proteins containing  $\gamma$ -tubulin (the  $\gamma$ TuRC) has been purified from *Xenopus* oocytes and shown to have MT-nucleating activity<sup>12</sup>. The *Xenopus* complex contains multiple copies of  $\gamma$ -

tubulin as well as at least 6 other proteins, of which only  $\alpha$ - and  $\beta$ -tubulin have been identified. Preliminary biochemical evidence suggests that a similar complex may exist in *Drosophila* (Y. Zheng, personal communication). The fact that the nucleating ring exists in organisms as divergent as *Xenopus* and *Drosophila* suggests that its complex structure is evolutionarily very old.

The similarity of the purified  $\gamma$ TuRC structure<sup>12</sup> to the ring structures seen in centrosomes suggests that MT-nucleating complexes are at least partially assembled in the cytoplasm before associating with the centrosome. Other centrosomal proteins presumably constitute the 'scaffolding' onto which the  $\gamma$ TuRC assembles at the centrosome. *In vivo*, the  $\gamma$ TuRC must become activated for MT nucleation only after it associates with the centrosome, because MTs do not appear randomly in the cytoplasm; perhaps, therefore, activation requires the binding of the ring structure to one or more PCM proteins. □

Received 29 August; accepted 9 November 1995.

1. Kellogg, D. R., Moritz, M. & Alberts, B. M. A. *Rev. Biochem.* **63**, 639–674 (1994).
2. Gould, R. R. & Borisy, G. G. *J. Cell Biol.* **73**, 601–615 (1977).
3. Oakley, B. R. *Trends Cell Biol.* **2**, 1–5 (1992).
4. Joshi, H. C., Palacios, M. J., McNamara, L. & Cleveland, D. W. *Nature* **356**, 80–83 (1992).
5. Felix, M.-A., Antony, C., Wright, M. & Maro, B. *J. Cell Biol.* **124**, 19–31 (1994).
6. Stearns, T. & Kirschner, M. *Cell* **76**, 623–637 (1994).
7. Moritz, M. et al. *J. Cell Biol.* **130**, 1149–1159 (1995).
8. Koster, A. J., Chen, H., Sedat, J. W. & Agard, D. A. *Ultramicroscopy* **46**, 207–227 (1992).
9. Koster, A. J. et al. *Microsc. Soc. Am. Bull.* **23**, 176–188 (1993).

10. Tilney, L. G. et al. *J. Cell Biol.* **59**, 267–275 (1973).
11. Evans, L., Mitchison, T. & Kirschner, M. J. *Cell Biol.* **100**, 1185–1191 (1985).
12. Zheng, Y., Wong, M. L., Alberts, B. & Mitchison, T. *Nature* **378**, 578–583 (1995).
13. Fung, J. C. et al. *J. struct. Biol.* (in the press).

**ACKNOWLEDGEMENTS.** We thank Y. Zheng and C. Field for antibodies, J. Fung and H. Chen for help with computer programs, M. L. Wong for help with controls, and Y. Zheng and T. Mitchison for discussions and for comments on the manuscript. M.M. was supported by postdoctoral fellowships from the American Cancer Society and the National Cancer Institute, and an NIH institutional training grant. This work was supported by the Howard Hughes Medical Institute (J.W.S. and D.A.A.), and by separate NIH grants to B.A., D.A.A. and J.W.S.



# Crystal structures of human calcineurin and the human FKBP12–FK506–calcineurin complex

Charles R. Kissinger, Hans E. Pargo, Daniel R. Knighton, Cristina T. Lewis, Laura A. Pelletier, Anna Tempczyk, Vincent J. Kalish, Kathleen D. Tucker, Richard E. Showalter, Ellen W. Moomaw, Louis N. Gastinel, Noriyuki Habuka, Xinghai Chen, Fausto Maldonado, John E. Barker, Russell Bacquet & J. Ernest Villafranca\*

Agouron Pharmaceuticals Inc., 3565 General Atomics Court, San Diego, California 92121-1121, USA

**CALCINEURIN (CaN)** is a calcium- and calmodulin-dependent protein serine/threonine phosphatase which is critical for several important cellular processes, including T-cell activation<sup>1</sup>. CaN is the target of the immunosuppressive drugs cyclosporin A and FK506, which inhibit CaN after forming complexes with cytoplasmic binding proteins (cyclophilin and FKBP12, respectively)<sup>2</sup>. We report here the crystal structures of full-length human CaN at 2.1 Å resolution and of the complex of human CaN with FKBP12–FK506 at 3.5 Å resolution. In the native CaN structure, an auto-inhibitory element binds at the Zn/Fe-containing active site. The metal-site geometry and active-site water structure suggest a catalytic mechanism involving nucleophilic attack on the substrate phosphate by a metal-activated water molecule. In the FKBP12–FK506–CaN complex, the auto-inhibitory element is displaced from the active site. The site of binding of FKBP12–FK506 appears to be shared by other non-competitive inhibitors of calcineurin, including a natural anchoring protein.

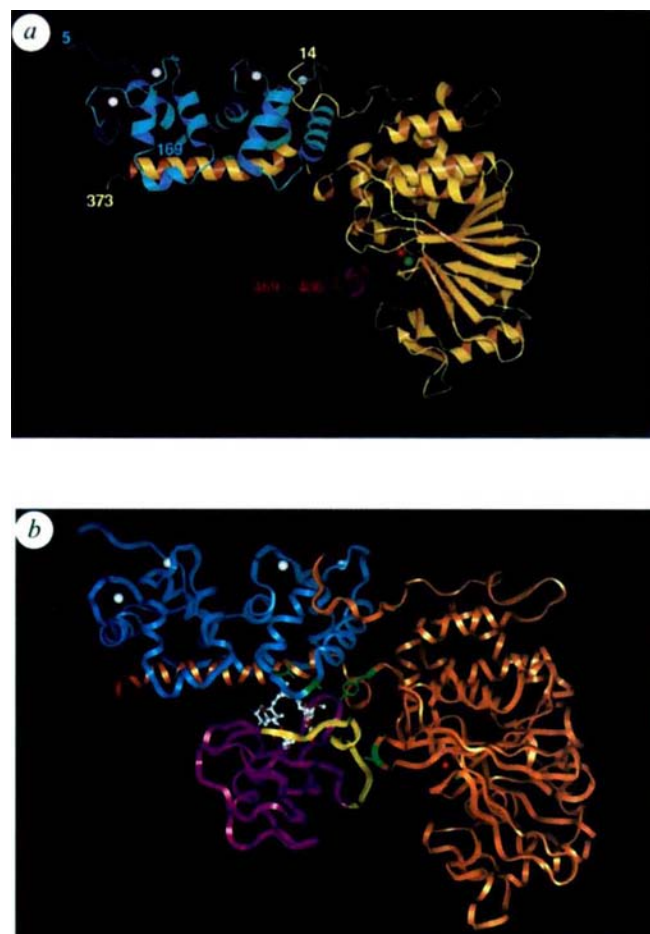
CaN is a heterodimer composed of an A subunit (CaNA; relative molecular mass 59K) and a B subunit (CaNB;  $M_r$  19K). Biochemical and genetic studies have identified four distinct functional domains of CaNA<sup>3–5</sup>: a catalytic domain, a CaNB-binding domain, a calmodulin (CaM)-binding domain, and an auto-inhibitory (AI) domain. CaNB binds four calcium ions and has 35% sequence identity with CaM<sup>6</sup>. CaN phosphatase activity is stimulated by  $\text{Ca}^{2+}$  binding to CaNB and by  $\text{Ca}^{2+}$ -induced binding of CaM to CaNA<sup>7</sup>.

We have determined the structure of the human CaN heterodimer at 2.1 Å resolution (Table 1 and Fig. 1a). CaNA contains a globular catalytic domain (residues 14–342) plus an extended, mostly  $\alpha$ -helical region (residues 343–373) which forms the CaNB-binding domain. CaNB, like CaM<sup>8</sup>, is composed of two lobes with two calcium ions bound by EF-hand motifs in each lobe. The overall architecture of these portions of the molecule appears to be substantially the same as described previously for a proteolytic fragment of bovine CaN bound to FKBP12–FK506 (ref. 9). A significant difference is that the amino-terminal segment of CaNA consisting of residues 14–23, which is not present in the fragment of the bovine enzyme, makes an important contribution to the CaNB-binding interface by forming part of a binding cleft for the carboxy-terminal lobe of CaNB (Fig. 1a). CaNA residues 1–13, 374–468 and 487–521 are not visible in the electron density map. These disordered regions include the CaM-binding domain predicted to lie between residues 390–414 (ref. 10).

An 18-residue segment of the C-terminal region of CaNA (Ser 469–Arg 486) lies over the apparent substrate-binding cleft

in the catalytic domain, comprising the bound portion of the AI domain (Figs 1a, 2). This AI segment consists of two short  $\alpha$ -helical regions plus five residues in an extended conformation. The residues that participate in binding interactions with the catalytic domain are the most highly conserved residues within the AI segment<sup>11</sup>. Residues on the C-terminal portion of the AI segment (Glu 481–Arg–Met–Pro 484) make the most extensive contacts with the substrate-binding cleft. Glu 481 hydrogen-bonds to water molecules bound to the two active-site metal ions. Although the helical regions of the AI segment contact the catalytic domain, their conformations preclude extensive interactions with the substrate-binding cleft (Fig. 2). The observed binding of the AI segment over the active site is consistent with two independent reports of competitive inhibition of CaN by a 25-residue peptide containing the AI segment using the RII phosphopeptide substrate from the regulatory subunit of cyclic AMP-dependent protein kinase<sup>12,13</sup>, although non-competitive inhibition with this substrate has also been reported<sup>14</sup>.

An intriguing question is the mechanism whereby calcium binding to CaNB stimulates CaN activity<sup>7</sup> in the absence of CaM. The C-terminal lobe, containing the higher-affinity cal-



**FIG. 1** a, Ribbon representation of the human CaN heterodimer showing secondary structural elements. The CaNA subunit is shown in yellow, the CaNB subunit in cyan. Fe (red) and Zn (green) ions are bound in a cleft in the catalytic domain. The AI segment is shown in red. Four calcium ions (white) are bound to CaNB. Both lobes of CaNB bind to the same side of the recognition helix. b, Ribbon representation of the FKBP12–FK506–CaN complex viewed as in a; CaNA is shown in orange and CaNB in blue. CaNA and CaNB residues implicated in cyclosporin A–cyclophilin binding (see text) are shown in green with side chains displayed. FKBP12 is shown in purple with the AKAP79 homology region of FKBP12 (see text) highlighted in yellow. FK506 is shown in ball-and-stick representation (C, white; N, blue; O, red). Fe and Zn in the CaNA active site are shown as red and green spheres, respectively.

\* To whom correspondence should be addressed.

FIG. 2 Stereo view of the AI segment of CaNA bound to the catalytic domain. The solvent-accessible surface (translucent grey) of the catalytic domain (yellow) is shown in the region of the substrate-binding cleft with the AI segment (C, green; N, blue; O, red) and bridging water molecules (pink spheres). The first and last residues of the AI segment visible in the electron density are labelled. Metal ions bound in the active site are shown as spheres (Zn, yellow; Fe, yellow and brown striped). The bound AI segment consists of the sequence Ser 469-Phe-Glu-Glu-Ala-Lys-Gly-Leu-Asp-Arg-Ile-Asn-Glu-Arg-Met-Pro-Pro-Arg 486. Ten residues of the AI sequence (470, 473, 474, 477, 480–485) and 20 residues of the catalytic domain form the binding interface. Seven hydrogen bonds are formed by main-chain or side-chain atoms of Asp 477, Glu 481, Arg 482, Met 483 and Pro 484; the rest of the interface consists mainly of hydrophobic interactions.

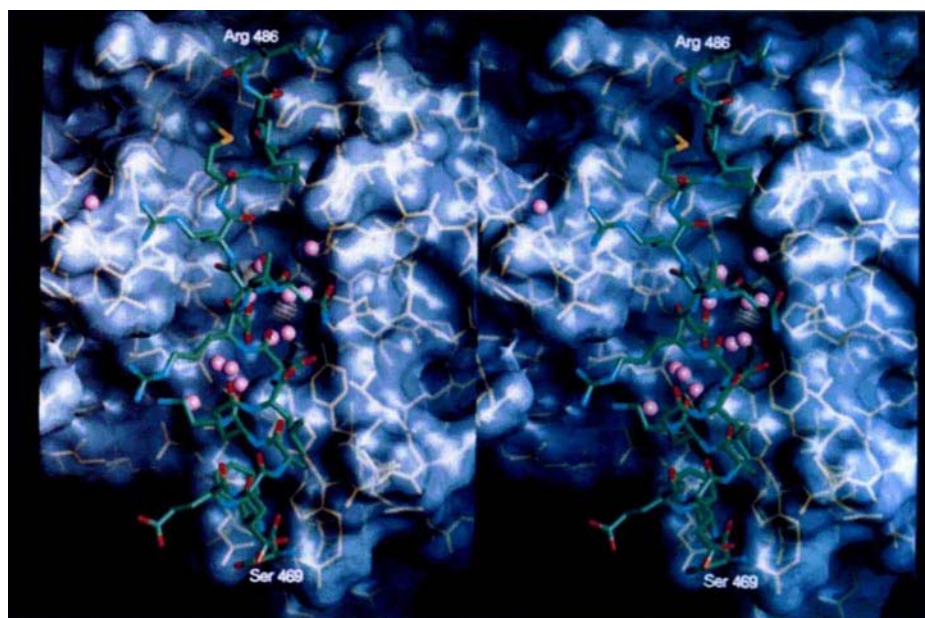


TABLE 1 Crystallographic structure determination

| Data set                | Resolution (Å) | Completeness (%) | Unique reflections | $R_{\text{sym}}^*$ (%) | $R_{\text{iso}}^\dagger$ (%) | Soak (mM/hours) | No. of sites | Phasing power $^\ddagger$ | $R_c^\S$ (%) |
|-------------------------|----------------|------------------|--------------------|------------------------|------------------------------|-----------------|--------------|---------------------------|--------------|
| CaN-native (1)          | 2.65           | 91.5             | 25,244             | 5.3                    |                              |                 |              |                           |              |
| CaN-native (2)          | 2.1            | 92.5             | 50,007             | 6.9                    |                              |                 |              |                           |              |
| CaN-PCMB                | 2.65           | 91.8             | 25,274             | 5.1                    | 8.4                          | 1/18            | 2            | 2.0/1.6                   | 56           |
| CaN-PIP                 | 3.0            | 97.9             | 18,557             | 8.6                    | 12.7                         | 0.1/4           | 1            | 0.84/0.61                 | 86           |
| FKBP12-FK506-CaN native | 3.5            | 96.3             | 11,495             | 10.6                   |                              |                 |              |                           |              |

Human recombinant CaN A and B subunits (CaNA $\alpha$ 1, CaNB $\beta$ 2) were coexpressed in *E. coli* and purified as described elsewhere<sup>13</sup>. The purified enzyme was homogeneous, full-length protein as evidenced by kinetic characteristics, SDS-PAGE and gel filtration<sup>13</sup>. Human CaN was crystallized from solutions containing polyethylene glycol (PEG) 6000 and CaCl<sub>2</sub> in space group  $P2_12_12_1$  ( $a=48.7$ ,  $b=104.3$ ,  $c=177.8$  Å, one molecule per asymmetric unit) as described elsewhere (L.A.P., et al., manuscript in preparation). The crystals were shown by SDS-PAGE to contain intact, non-proteolysed CaN. The FKBP12-FK506-CaN complex was crystallized at room temperature by vapour diffusion in hanging drops (space group  $P4_32_12$ ,  $a=b=104.9$  Å and  $c=162.6$  Å, one complex in the asymmetric unit). The reservoir solutions contained 6% PEG 6000, 10 mM HoCl<sub>3</sub>, 0.1 M bis-Tris, pH 6.8, and 1 mM DTT. X-ray diffraction data were collected at the Cornell High Energy Synchrotron Source using phosphor image plates and at the European Synchrotron Radiation Facility (Grenoble, France) using a MAR image plate system ( $\lambda=0.908$  Å). All data sets were collected under a stream of cold nitrogen ( $-181$  °C) from single crystals flash-cooled in 30% glycerol. Data were converted to integrated intensities using DENZO and SCALEPACK (Z. Otwinowski). The structure of human CaN was determined by multiple isomorphous replacement plus anomalous scattering (MIRAS) phasing and refined using X-PLOR 3.1 (ref. 26). An anisotropic  $B$  factor correction was applied once using X-PLOR. The final refined model contains CaNA residues 14–373 and 469–486, CaNB residues 5–169, 4 calcium ions, one zinc, one iron and 436 water molecules. The mean  $B$  factor of the model is 39.2 Å<sup>2</sup>. The final  $R$  factor is 0.187 for 48,504 reflections ( $F>0$ ) in the resolution range 10–2.1 Å (90.8% of possible data). The root-mean-square deviations of bond lengths and bond angles from ideal values are 0.019 Å and 1.94°, respectively. Three residues in CaNA that participate in metal binding (Asp 121, Arg 122, His 281) are in disallowed regions of a Ramachandran plot as defined by PROCHECK<sup>27</sup>. The structure of the FKBP12-FK506-CaN complex was determined by molecular replacement (MR) using the coordinates of the human CaN heterodimer as a search model. MR calculations were carried out using the program AMoRe<sup>28</sup> in the CCP4 program system. The initial MR solution gave an  $R$  factor of 0.535 and correlation coefficient of 0.162 for all data in the resolution range 8–4 Å. After subjecting the CaN model to rigid-body refinement, it was possible to identify the FKBP12-FK506 in a difference electron-density map. The coordinates of the FKBP12-FK506 complex determined at 1.7 Å resolution<sup>29</sup> were fitted to the electron density. Repetition of the MR calculations using the combined FKBP12-FK506-CaN model produced the same solution with an  $R$  factor of 0.451 and correlation coefficient of 0.444. Refinement of the MR solution as a single rigid-body in AMoRe<sup>28</sup> resulted in an  $R$  factor of 0.395 for all data in the resolution range 20–4.0 Å. Refinement was completed using X-PLOR. Because of the limited quantity of data, careful attention was paid to the behaviour of the free  $R$  factor<sup>26</sup> throughout the course of the refinement. The free  $R$  factor was calculated using 10% of the data omitted from the refinement calculations. Refinement with CaN and FKBP12-FK506 as separate rigid-bodies followed by positional refinement of atomic parameters resulted in an  $R$  factor of 0.262 for data in the resolution range 10–3.5 Å (free  $R$  factor, 0.349) with a single overall temperature factor. The final model includes residues 14–374 of CaNA, residues 5–169 of CaNB, residues 1–107 of FKBP12 and FK506.

\*  $\sum_h \sum_i |I_{h,i} - \langle I_h \rangle| / \sum_h \sum_i I_{h,i}$  where  $\langle I_h \rangle$  is the mean intensity for the  $i$  observations of reflection  $h$ .

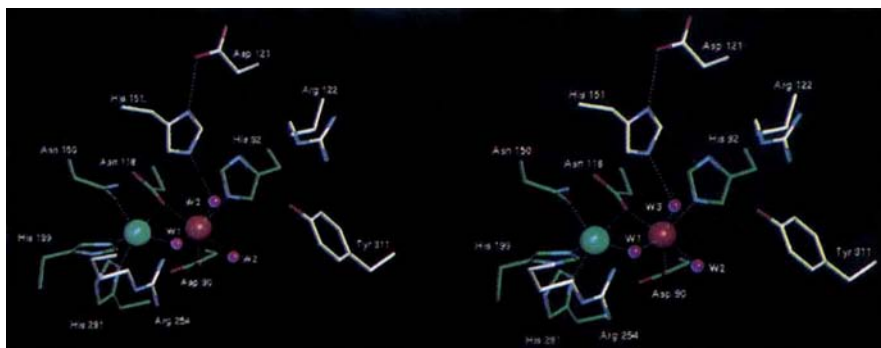
$^\dagger \sum |F_{\text{der}} - F_{\text{nat}}| / \sum F_{\text{nat}}$  where  $F_{\text{der}}$  and  $F_{\text{nat}}$  are the derivative and native structure factor amplitudes, respectively.

$^\ddagger [F_{\text{H(calc)}}]^2 / ([F_{\text{PH(obs)}}] - [F_{\text{PH(calc)}}])^{1/2}$ , given for acentric/centric reflections.

$^\S \sum |F_{\text{der}} \pm F_{\text{nat}}| - F_{\text{H(calc)}} / \sum |F_{\text{der}} - F_{\text{nat}}|$  for centric reflections, where  $F_{\text{H(calc)}}$  is the calculated heavy-atom structure factor.



FIG. 3 Stereo view of CaN active site. Metal-ligating side chains (C, green; N, blue; O, red) and conserved active-site side chains (C, white; N, blue; O, red) are shown. The zinc (green) and iron (red) are 3.14 Å apart. The iron is octahedrally coordinated by Asp 90 Oδ1 (2.31 Å), His 92 Nε2 (2.37 Å), Asp 118 Oδ1 (2.26 Å), water W2 (2.50 Å), water W1 (2.66 Å), and water W3 (2.79 Å). The zinc coordination is a distorted trigonal bipyramid formed by Asp 118 Oδ1 (2.60 Å), Asn 150 Oδ1 (2.11 Å), His 199 Nε2 (2.30 Å), His 281 Nδ1 (2.29 Å), and water W1 (2.42 Å). W3 is bound to the iron and is also hydrogen-bonded to Nε2 of His 151, a conserved residue that does not participate in metal binding but is poised over the dimetal site. Two additional conserved residues, Arg 122 and Arg 254, are well positioned to participate in binding to the phosphate oxygens of the substrate.



cium sites, is the most intimately associated with CaNA. This lobe may serve to anchor the subunits and transduce the effects of calcium binding at the lower-affinity (micromolar) sites<sup>7,15</sup> on the N-terminal lobe to the catalytic domain.

We assigned the positions of the iron and zinc at the active site of human CaN (Fig. 3) on the basis of the similarity of the dimetal site to the zinc-iron site in the structure of kidney bean purple acid phosphatase<sup>16</sup> (KBPA). The same assignments were made for the dimetal site in the structure of bovine CaN<sup>9</sup>. Spectroscopic studies have given conflicting evidence as to whether the bound iron exists in the Fe<sup>2+</sup> or Fe<sup>3+</sup> form in the catalytically active enzyme<sup>17,18</sup>. Several independent kinetic and biochemical experiments, primarily using small phosphotyrosine substrates, indicate that direct hydrolysis of the phosphoester bond by CaN occurs without generation of a phosphoryl-enzyme intermediate<sup>6,19</sup>. There are three metal-bound water molecules at the active site that could potentially represent the attacking nucleophile (Fig. 3). Two of these, a metal-bridging water molecule (W1) and a water molecule bound by the iron (W2), occupy positions similar to those of water molecules observed at the dimetal site of protein serine/threonine phosphatase-1 (PP-1)<sup>20</sup>. A third water molecule (W3) is bound by the iron and hydrogen bonded to His 151 (Fig. 3). Binding of a water molecule at this position in the PP-1 structure appears to be sterically precluded by the bound inhibitor, microcystin. The His 151 side chain forms a hydrogen bond with the side chain of Asp 121 (Fig. 3) and is thus a potential proton acceptor from W3. Our modelling experiments indicate that an activated water molecule in this position could be the attacking nucleophile. An iron-bound hydroxide ion is also postulated to be the attacking nucleophile in KBPA<sup>16</sup>.

The structure of the FKBP12-FK506-CaN complex was determined at 3.5 Å resolution (Table 1, Fig. 1b). Neither CaN nor FKBP12 undergoes structural changes upon complex formation that are detectable at this resolution. As in the crystal structure of native CaN, the C-terminal portion of CaNA is disordered beyond the CaNB-binding helix, and electron density for the CaM-binding domain of CaNA is not visible. Unlike the native CaN crystal structure, the AI segment of CaN is not seen at the CaN active site in the complex nor elsewhere in the crystal structure. FKBP12-FK506 contacts two distinct areas on CaN (Fig. 1b). At the large site of interaction (primary recognition site), the composite surface formed by FK506 and surrounding residues of FKBP12 contacts residues on the exposed side of the CaNB-binding helix of CaNA and surrounding residues from

both lobes of CaNB (Fig. 1b). Additional interactions are seen between FKBP12 and the catalytic domain of CaNA near one side of the substrate binding cleft (secondary recognition site). The binding interactions appear to correspond closely to those seen in the complex between FKBP12-FK506 and a proteolytic fragment of bovine CaN<sup>9</sup>.

We have found that FKBP12-FK506 exhibits classical non-competitive inhibition of CaN with the RII peptide and with the protein substrate, DARPP-32 (C.T.L., manuscript in preparation). Non-competitive inhibition generally implies formation of a catalytically inactive enzyme-substrate-inhibitor complex. The observed binding of FKBP12-FK506 to CaN is consistent with formation of such a complex in that no part of the active site or substrate-binding cleft of CaN is directly obstructed. Our modelling studies indicate that FKBP12-FK506 would not sterically prevent binding of peptide substrates or the AI segment at the CaNA active site. Comparison of the structures of CaN and the FKBP12-FK506-CaN complex indicates that large conformational changes in CaN are not responsible for inhibition by FKBP12-FK506 (Fig. 1a, b). However, contact by FKBP12-FK506, particularly at the secondary recognition site (residues 310-314 and 159 of CaNA), could produce subtle changes in active-site geometry, not detectable at the resolution of this study, that result in inhibition. An alteration of active-site geometry might also explain the interesting observation that binding of FKBP12-FK506 or cyclophilin-cyclosporin A (CyP-CsA) results in a slight increase in CaN phosphatase activity when *p*-nitrophenyl phosphate is used as the substrate instead of a phosphopeptide<sup>2,21</sup>.

The residues in CaN identified by genetic methods to be important in CyP-CsA binding<sup>22,23</sup> are found at the primary or secondary recognition sites for FKBP12-FK506 (Fig. 1b). CyP-CsA, like FKBP12-FK506, is a non-competitive inhibitor of CaN when assayed with a phosphopeptide substrate<sup>24</sup>. It appears, therefore, that CyP-CsA binds at the same general location as FKBP12-FK506 and that it may inhibit CaN through a similar mechanism. An anchor protein, AKAP79 (A-kinase anchoring protein 79), is also a non-competitive inhibitor of CaN<sup>25</sup>. The putative CaN-binding domain of AKAP79 contains a region with a high degree of sequence similarity to residues 32-47 of FKBP12. A 22-residue constituent peptide encompassing this FKBP12 homology region also inhibits CaN<sup>25</sup>. In the FKBP12-FK506-CaN complex, the corresponding FKBP12 residues are involved extensively in interactions with CaN (Fig. 1b). These residues constitute the entire surface of FKBP12 con-

tacting the secondary recognition site, and a large portion of the surface contacting the primary recognition site. Therefore, AKAP79 also appears to bind at the same site as FKBP12–FK506 and may inhibit CaN by a similar mechanism. The observation that CaN is inhibited by a 22-residue peptide that would appear to provide little direct steric hindrance to substrate binding at the active site provides additional support for the suggestion that factors other than steric hindrance contribute to CaN inhibition. □

Received 30 August; accepted 3 November 1995.

- Shenolikar, S. & Nairn, A. C. in *Advances in Second Messenger and Phosphoprotein Research* (eds Greengard, P. & Robison, G. A.) 1–121 (Raven, New York, 1991).
- Liu, J. et al. *Cell* **66**, 807–815 (1991).
- Klee, C. B., Crouch, T. H. & Krinks, M. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6270–6273 (1979).
- Hashimoto, Y., Perrino, B. A. & Soderling, T. R. *J. biol. Chem.* **265**, 1924–1927 (1990).
- Watanabe, Y., Perrino, B. A., Chang, B. H. & Soderling, T. R. *J. biol. Chem.* **270**, 456–460 (1995).
- Klee, C. B., Draetta, G. F. & Hubbard, M. J. *Adv. Enzym. Relat. Areas molec. Biol.* **61**, 149–200 (1988).
- Stemmer, P. M. & Klee, C. B. *Biochemistry* **33**, 6859–6866 (1994).
- Babu, Y. S., Bugg, C. E. & Cook, W. J. *J. molec. Biol.* **203**, 191–204 (1988).
- Griffith, J. P. et al. *Cell* **82**, 507–522 (1995).
- Perrino, B. A., Ng, L. Y. & Soderling, T. R. *J. biol. Chem.* **270**, 340–346 (1995).
- Guerini, D., Montell, C. & Klee, C. B. *J. biol. Chem.* **267**, 22542–22549 (1992).
- Parsons, J. N. et al. *J. biol. Chem.* **269**, 19610–19616 (1994).
- Lewis, C. et al. *FASEB J.* **9**, A1346 (1995).
- Perrino, B. A., Hashimoto, Y. & Soderling, T. R. *FASEB J.* **4**, A2237 (1990).
- Burroughs, S. E., Horrocks, W. D. Jr, Ren, H. & Klee, C. B. *Biochemistry* **33**, 10428–10436 (1994).
- Sträter, N., Klambunde, T., Tucker, P., Witzel, H. & Krebs, B. *Science* **268**, 1489–1492 (1995).
- Yu, L., Haddy, A. & Rusnak, F. J. *Am. chem. Soc.* **117**, 10147–10148 (1995).
- Qin, K.-F., Khangulov, S., Liu, C. & Huang, C. Y. *FASEB J.* **9**, A1347 (1995).
- Martin, B. L. & Graves, D. J. *Biochim. biophys. Acta* **1206**, 136–142 (1994).
- Goldberg, J. et al. *Nature* **376**, 745–753 (1995).
- Swanson, S. K. H. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 3741–3745 (1992).
- Cardenas, M. E., Muir, R. S., Breuder, T. & Joseph, H. *EMBO J.* **14**, 2772–2783 (1995).
- Milan, D., Griffith, J., Su, M., Price, E. R. & McKeon, F. *Cell* **79**, 437–447 (1994).
- Etzkorn, F. A., Chang, Z., Stolz, L. A. & Walsh, C. T. *Biochemistry* **33**, 2380–2388 (1994).
- Coghlan, V. M. et al. *Science* **267**, 108–111 (1995).
- Brünger, A. T. *X-PLOR Version 3.1 Manual* (Yale Univ., New Haven, CT, 1993).
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. *J. appl. crystallogr.* **26**, 283–291 (1993).
- Navaza, J. *Acta crystallogr.* **A50**, 157–163 (1994).
- Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. *Science* **252**, 839–842 (1991).

**ACKNOWLEDGEMENTS.** We dedicate this publication to the memory of Dr Russell Bacquet. We thank R. Kincaid for CaN $\alpha$  cDNA; C. Klee for CaN $\beta$  and CaN $\gamma$  cDNAs; S. Ealick for data collection time at CHESS and R. Walter for assistance with data collection; the fermentation team at Agouron Pharmaceuticals (J. Dutra, H. Tenenbaum, Z. Hostomsky and S. Rahmati) for large quantities of *E. coli* coexpressing human CaN; our colleagues X. Hou, R. Babine, P. Dragovich, K. Menge and W. Chong for helpful discussions, and S. Fordyce for assistance in preparation of this manuscript. We also acknowledge the support of our colleagues in the immunosuppression program at Agouron Pharmaceuticals and our collaborators at JT Pharmaceuticals.

## ERRATA

### A protein catalytic framework with an N-terminal nucleophile is capable of self-activation

**James A. Brannigan, Guy Dodson, Helen J. Duggleby, Peter C. E. Moody, Janet L. Smith, Diana R. Tomchick & Alexey G. Murzin**

*Nature* **378**, 416–419 (1995)

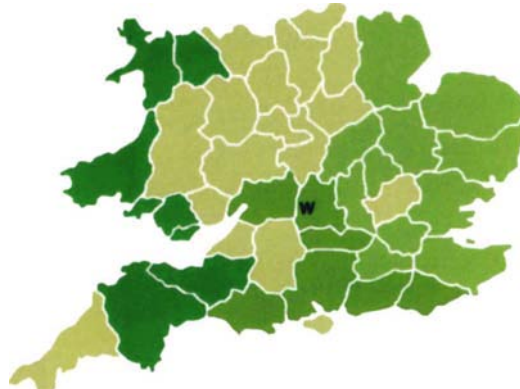
As a result of an error during the production process, Fig. 4 of this Letter was printed upside down. □

### Predation risk and the cost of being fat

**Andrew G. Gosler, Jeremy J. D. Greenwood & Christopher Perrins**

*Nature* **377**, 621–623 (1995)

THE map in Fig. 2 should have been marked with a 'W', as indicated here, to represent Wytham. Counties with insufficient data are shown in beige, and not in white as stated in the figure legend. □

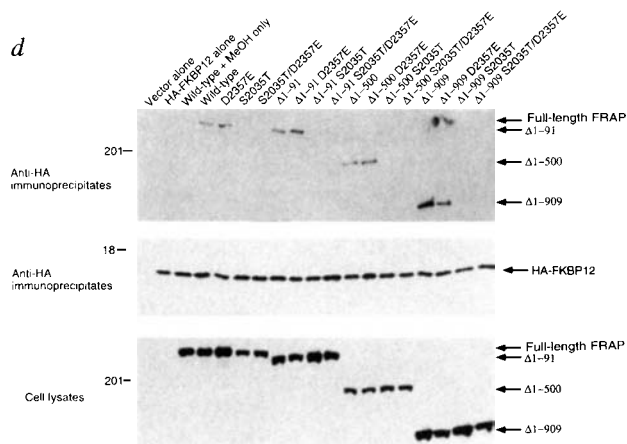


### Control of p70 S6 kinase by kinase activity of FRAP *in vivo*

**Eric J. Brown, Peter A. Beal, Curtis T. Keith, Jie Chen, Tae Bum Shin & Stuart L. Schreiber**

*Nature* **377**, 441–446 (1995)

THE middle panel of Fig. 4d of this Letter did not print up properly. The complete figure is shown here. □



# Non-invasive ion probes — tools for measuring transmembrane ion flux

Peter J. S. Smith

**Ionophores, used in non-invasive vibrating electrodes, can directly monitor steady-state ion flux from single cells, even when the source is non-electrogenic.**

LIVING organisms have a capacity to regulate their internal ionic environment at compositions very different from the external milieu and to expend energy to achieve this. The electrochemical gradients established are central to many facultative and specialized cell functions, with some of the membrane-borne ion transport mechanisms lending themselves to study with established electrophysiological techniques, such as voltage or current clamp and microelectrodes. However, in the case of the relatively steady-state ion exchange between the cell and its environment the very small associated net voltage changes or non-electrogenic events generated by the activity of ATPases and exchangers are seldom directly detectable by conventional means, particularly at the single-cell level. If a method could be devised to observe the flux of a single ion species across the plasma membrane, then these important aspects of ion homeostasis could be brought within the realm of the cellular physiologist. These are the problems solved by the vibrating voltage and ion-selective probes, instruments developed at the National Vibrating Probe Facility of the National Institutes of Health (National Center for Research Resources).

## Background to vibrating probes

The non-invasive vibrating ion-selective probe (NVP<sub>i</sub>, refs 1, 2) is based on the same basic principles as the earlier non-invasive vibrating voltage probe (NVP<sub>PD</sub>, refs 3, 4). Both measure the weak extracellular voltages associated with relatively steady-state ion flux in a conductive medium surrounding cells. Measuring these weak voltages is a challenging problem. One could attempt this by placing two glass electrodes, filled with liquid electrolyte, at a known distance from the cell and a known distance apart. By using the continuous version of Ohm's law the current passing between the two electrodes could be calculated from the procedure. An alternative solution, to vibrate a single, and therefore self-referencing, electrode, has revolutionized the study of electric fields associated with living tissue<sup>1</sup>.

Jaffe and Nuccitelli solved several problems in the measurement of the

transmembrane ion flux with the introduction of the high capacitance (~10 nF) platinum-black voltage probe<sup>1</sup>. This low-noise metal electrode drastically reduced the interface impedance, which reduced thermal noise, so that a vibrational frequency of a few hundred hertz could extract potential differences of a few nanovolts. Standard lock-in amplifiers could be used to discriminate the signal of known frequency from the background noise. The technique was made all the more powerful by the introduction of a computer-based, two-dimensional system, allowing the calculation of current vectors with a degree of directionality<sup>5</sup>. With this device it was possible to map fields in a wide variety of situations.

Powerful and sensitive though the NVP<sub>PD</sub> is, it has several limitations. The size of the platinum black ball (5–10 µm) and the frequency of vibration (~200 Hz) creates a vortex-like mixing at the probe tip, mechanically disturbing all but quite robust preparations. Furthermore, the measurement of net voltages offers limited information on the ions carrying the underlying current. The newer ion-selective vibrating probe has proved to be a promising approach to the measurement of transmembrane ion flux, minimizing some of the shortcomings of the original technique.

## Ion-selective vibrating probe

Although different from the voltage probe, the ion-selective probe depends on the same basic principle of a vibrating self-referencing electrode. Here, however, the objective is to measure the net flux of a single ion species across the plasma membrane of a single cell or defined tissue region. To achieve this we take advantage of commercially available ionophores. Most of these ionophores are

macrocyclic neutral ion carriers suspended in a membrane solvent. These carriers are lipophilic complexing agents capable of transporting ions selectively across organic membranes, probably by a carrier relay mechanism<sup>6</sup>. These ionophores have been used in stationary intracellular electrodes where difficulties can arise from competing ions, voltage sensitivities and electrode drift. In the applica-

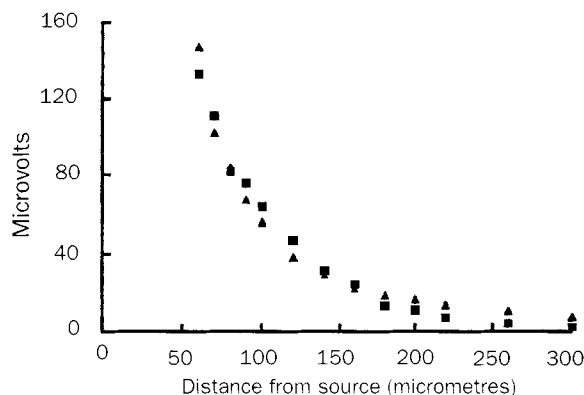


FIG. 1 An important step in the development of the ion-selective probe was to demonstrate that voltage differences measured in a steady-state gradient complied with a theoretical model of the same gradient. This figure illustrates such an experiment where the theoretical gradient (▲) is calculated as  $\Delta V = S[(-K\Delta r)(C_0r^2 + Kr)^{-1}](2.3)^{-1}$  where  $\Delta V$  is the expected potential difference between two points  $\Delta r$  apart (10 mm in this example) at a distance of  $r$  from an artificial source, defined by the empirical constant  $K^{1,2}$ .  $S$  is the Nernst slope of the electrode. The measured values (■) comply well with the expected values. Flux values can be calculated by substituting the  $\Delta\mu V$  values into the equation  $\Delta C \approx [2.3(\Delta VC_B)]S^{-1}$  where  $C_B$  is the background concentration of calcium in the medium. Substitution of  $\Delta C$  into the Fick equation yields a flux<sup>2</sup>.

tions described here, however, the self-referencing principle reduces the impact of drift and extracts only differential signals. Furthermore, being extracellular, the probe is normally located in the weakest of electric fields.

Non-invasive vibrating ion-selective electrodes have their own set of requirements and limitations, most of which are connected with the need to move the electrode at frequencies below 1 Hz, over a 10- to 50-µm distance, in close proximity to a cell. A balance must be achieved between a frequency that reduces the electrode drift but allows the establish-

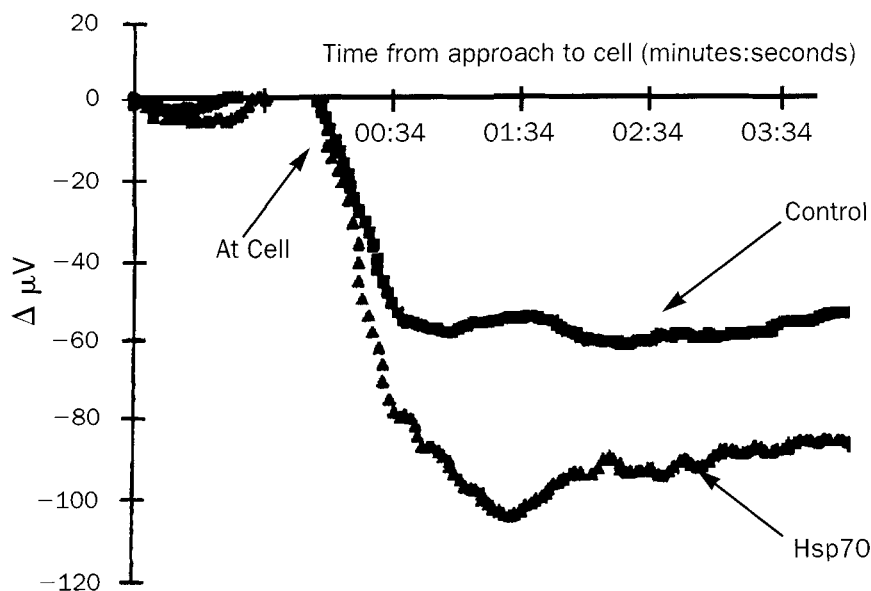


FIG. 2 Calcium flux data are acquired from single cells by placing the ion-selective electrode in close proximity to the plasma membrane and moving it over 10  $\mu\text{m}$  at 0.3 Hz (■). As the probe collects data at the cell surface the averaging buffer fills with new values and a decreasing  $\Delta\mu\text{V}$  reading is recorded from the previous stable background position located away from the cell (▲). By convention, a negative value represents a recordable efflux from the cell. These fluxes can be modified by pharmacological agents, one such being a heat shock protein of 70K size<sup>9</sup>.

ment of the new ion concentration at the electrode tip<sup>1</sup>. For calcium ionophore 1 cocktail A (Fluka — based on the neutral carrier ETH1001), this frequency is around 0.3 Hz. Hydrogen ions, with faster diffusion, re-establish gradients more rapidly and can be measured at higher frequencies (0.1 Hz — Fluka Hydrogen ionophore 1 (triiododecylamine) cocktail A), although corrections for buffer-mediated diffusion must be made<sup>7</sup>. In both cases there is an unavoidable attenuation of the signal based on the response time of the ionophore (90 per cent, < 5 s for both calcium and hydrogen mentioned above), which results in an inaccurate measurement of weak gradients in relation to the theoretical values expected (Fig. 1). In general, however, the technique is remarkably sensitive having, for calcium measurements, a noise level of around 2–8  $\mu\text{V}$ , and a sensitivity of less than 1  $\text{pmol cm}^{-2} \text{sec}^{-1}$  (equivalent to a current density of 0.2  $\mu\text{A cm}^{-2}$ ). The sensitivity is acquired by extensive signal sampling and averaging.

Electrode construction and the electronics differ little from conventional ion-selective electrode techniques. Glass electrodes are pulled as for patch electrodes, silanized, back-filled with electrolyte and front-filled with a short column of ionophore<sup>2</sup>. The head-stage contains a unity gain field effect operational amplifier (input impedance  $10^{15} \Omega$ ) configured as a voltage follower. Signals from excursion extremes are passed through high- and low-pass filters, amplified by 1,000 times and passed to the A-D board of a 486PC at a sampling rate of

1,000 Hz. Signals from each position are packaged into ten groups, the first three after movement being discarded. Differential values in microvolts are calculated from the two positions and fed into a running average. The controlling software is custom-made to allow both data collection and control of all vibrational parameters, including excursion distance and angle of attack. Electrodes are calibrated before experiments to confirm the Nernstian response. Thus, given known distance of vibration, background concentrations of the measured ions, diffusion constants and electrode slopes, microvolt signals can be converted to flux data and current densities<sup>2</sup>. Although clearly dependent on the signal-to-noise ratio, it should be obvious from the above that this technique will not record fast transient events, focusing instead on activities that last for several seconds to tens of seconds.

## Applications

Achieving the accurate and stable vibration of the electrodes is a critical aspect of the ion-selective technique, particularly when considering the movement of ions across the plasma membrane of a single cell. The original design, based on an orthogonal array of piezoelectric motors<sup>1</sup>, has been replaced by micro-stepper motors controlled via the computer parallel port<sup>2</sup>. This arrangement gives the experimenter more control over electrode position and movement. A controlled acceleration profile is particularly important in providing smooth motion when in close proximity to a cell.

By using this method the probe can approach to within 2  $\mu\text{m}$  of the plasma membrane of isolated *Aplysia* (sea snail) bag cell neurons, a test preparation that has proven particularly useful in the development of this technique. For studying transmembrane steady-state calcium flux it has proven ideal (Fig. 2). These cells exhibit a modulated calcium efflux<sup>8,9</sup> from the soma, which persists over several days in culture. Similar fluxes have also been recorded from several other isolated cell preparations<sup>2,10,11</sup>.

## Summary

Although originally developed to study transmembrane calcium flux, the vibrating ion-selective probes are not restricted to this ion or to single-cell applications. To date, ionophores for potassium<sup>12,13</sup> and hydrogen<sup>12</sup> have been successfully used, with development studies completed for magnesium and chloride (J. G. Kunkel *et al.* manuscript in preparation). The capability of this approach for selecting one ion in a co-transporting pair, as in the H1/K1ATPase and  $\text{Na}^+/\text{Ca}^{2+}$  antiporter, or the ubiquitous, but electrically neutral  $\text{Na}^+/\text{K}^+/\text{2Cl}^-$  symporter, makes it a powerful technique. The further development and application of these techniques are the goals of the National Vibrating Probe Facility. As a national and international facility of the National Institutes of Health, we encourage scientists to apply for rig time in the facility, providing laboratory space and technical assistance. □

Peter J. S. Smith is an associate scientist at the Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA, and director of the National Vibrating Probe Facility, funded through the NIH's National Center for Research Resources. For more information, fill in reader service number 100 on the reader service card bound inside the journal.

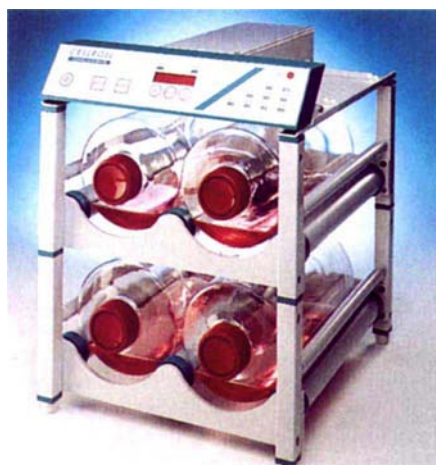
- Kuhtreiber, W. M. & Jaffe, L. F. *J. Cell Biol.* **110**, 1565–1573 (1990).
- Smith, P. J. S., Sanger, R. H. & Jaffe, L. F. *Meth. Cell Biol.* **40**, 115–134 (1994).
- Jaffe, L. F. & Nuccitelli, R. *J. Cell Biol.* **63**, 614–628 (1976).
- Nuccitelli, R. In *Noninvasive Techniques in Cell Biology*, (Wiley-Liss, New York, 1990).
- Scheffey, C. *Rev. Sci. Instr.* **59**, 787–792 (1988).
- Ammann, D. *Ion-Selective Microelectrodes: Principles, Design and Application*. (Springer-Verlag, New York, 1986).
- Demarest, J. R. & Morgan, J. L. M. *Biol. Bull.* **189**, (in the press).
- Yamashita, E. & Smith, P. J. S. *Biol. Bull.* **187**, 270 (1994).
- Smith, P. J. S., Hammar, K. & Tytell, M. *Biol. Bull.* **189**, (in the press).
- Keefe, D., Pepperell, J., Rinaudo, P., Kunkel, J. & Smith, P. *Biol. Bull.* **189**, (in the press).
- Land, S. C. & Smith, P. J. S. *Biol. Bull.* **189**, (in the press).
- Kochian, L. V., Shaff, J. E., Kuhtreiber, W. M., Jaffe, L. F. & Lucas, W. J. *Planta* **188**, 601–610 (1992).
- Marcus, D. C. & Shipley, A. M. *Biophysical J.* **66**, 1939–1942 (1994).



# Cell studies

New research tools for the cell biologist include a fermenter, serum-free medium for insect cell lines, inverted research microscopes, and intracellular ion measurement and digital video microscopy systems.

CELLROLL and Cellspin from Integra Biosciences combine to offer scientists the flexibility to both produce routine cell products on a large scale and successfully **cultivate problematic cell lines** (*Reader Service No. 101*). Their compact design, low heat generation and humidity resistance allows them to be operated remotely through a universal control panel within a CO<sub>2</sub> incubator. Cellroll allows up to eight decks of two roller bottles each to be controlled from a single



Cellroll from Integra Biosciences.

drive unit. Stacking the decks in different ways enables optimal utilization of space within CO<sub>2</sub> incubators and warm rooms. Both standard velocity, 0.1- to 2-r.p.m., and high-velocity 2- to 6-r.p.m. drive units are available. To tackle the problem of mixing cells vigorously enough to form a perfect suspension yet gently enough to avoid damage, Cellspin incorporates a unique pendular system with a transverse oscillatory action and increased surface area-to-volume ratio. The sophisticated control panel allows continuous stirring at speeds between 5 and 100 r.p.m. Cellspin's compact construction means that up to four mixing units can be treated in one 216-litre incubator. The universal smart control panel recognizes Cellspin and Cellroll drive units.

Ex-Cell 405 **serum-free medium for insect cell lines** is now available from JRH Biosciences to provide a serum-free environment for optimal growth and viability for suspended BTI-TN-5B 1-4 cells (High Five) (*Reader Service No. 102*). In addition, protein expression is not compromised through the transition to suspension and the incorporation of serum-free medium. The medium is said to support

recombinant protein expression from BTI-TN-5B 1-4 cells to levels equal to or greater than seen in other media.

A new density perturbation technique using Accu-beads from Accurate Chemical and Scientific now allows for the easy separation of cells on the basis of the immunological identity of the cell surface with the aid of **antibody-coated, non-magnetic beads** (*Reader Service No. 103*). The use of antibody-coated beads to sort cells is a well established method. Similarly, the use of magnetic beads is well documented. However, these beads are very expensive. The Accu-beads are significantly less expensive and can be used to give similar separations. The less expensive non-magnetic density perturbation can be used to substitute existing immunomagnetic separations such as those for negative selection, for the elimination of a specific population. This method should prove useful for the separation of subsets of tissue culture cells and peripheral blood cells.

Bioengineering has designed a new **fermenter** to help maintain sterile continuous ultrafiltration of fermenter broth (*Reader Service No. 104*). This development enables the integration of ultrafiltration devices in a fermenter and reduces product inhibition in the fermenter, where the product is formed. The modular unit is integrated in a standard fermenter. Filtration capacity is flexible as the number of filtration disks is variable. The installation of blade stirrers in between the filtration disks minimizes fouling on the membranes. This approach is designed to eliminate the sterilization or autoclaving of bypass lines, the circulation pump and the filter module.

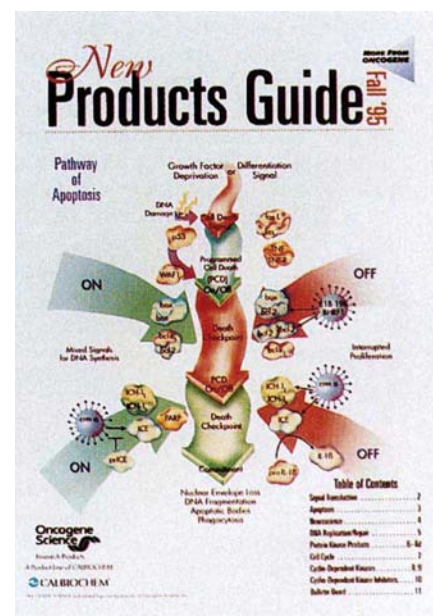
Becton Dickinson Immunocytometry Systems has announced the availability of ProCount, the first research reagent system for the **enumeration of CD34<sup>+</sup> haematopoietic progenitor cells** from peripheral blood (*Reader Service No. 105*). The easy-to-use assay for flow cytometry, available for research use only, features a combination of highly specific monoclonal antibodies, a nucleic-acid dye and absolute counting beads. Besides yielding absolute counts of CD34 cells in peripheral blood or leukapheresis samples, the assay also determines CD34 cells as a percentage of both total leucocytes and of total nucleated cells. The method owes its reliability to advanced flow cytometry technology, com-



Becton Dickinson's ProCount system.

bined with a unique reagent formulation. The ProCount reagent assay has been designed to work with standard acquisition and analysis software.

Oncogene Science/Calbiochem now offers its autumn 1995 **new products guide** featuring more than 200 new products (*Reader Service No. 106*). Thirty new antibodies, including four antibodies recognizing proteins involved in apoptosis; fas, bax, bcl-2, and bcl-x are offered. Over 160 protein kinase C products are featured,



Autumn 1995 new products guide.

along with a peptide-specific antisera for neuroscience research, which is screened for the ability to stain floating rat brain tissue sections. Antibodies to proteins involved in signal transduction, cell cycle, neuroscience, DNA replication and repair, are also offered.

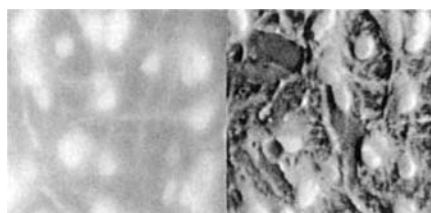
## Under the microscope

The IX70, a new modular **inverted research microscope** from Olympus is designed to provide ease of use and versatility for research applications (*Reader Service No. 107*). The U.I.S. optics of this microscope have all corrections completed within the objectives. Performance in each microscopy mode is optimized, including bright-field, phase-contrast, Nomarski DIC and incident-light fluorescence. For all modes, the F.N. 22 extended field of view covers 21 per cent more area than the F.N. 20 commonly used. It has a 100-W halogen, transmitted-light illumination pillar that tilts 30 degrees and includes a rotatable centre section, which has the ability to accept two back-to-back condensers. A built-in power supply in the left wing of the stand is isolated from all optical components. When the microscope is used in a Faraday cage for electrophysiology work, an external 12 V/100 W power supply can be kept outside the cage. A left-side photo/video port and a front-mounted 35-mm SLR camera or video port are standard. Access to the right side port and direct bottom port is available as an option. Every photo/video port can be used as a light input axis, and all kinds of laser adaptation can be easily performed.

Carl Zeiss has introduced a new **scanning electron microscope**, called the DSM 982 Gemini, which is said to be capable of achieving a lateral resolution as low as 3 nm at a beam energy of 1 keV, and 1 nm at 20 keV (*Reader Service No. 108*). The new low-voltage SEM adopts a new design concept for the electron optical column, with a compound magnetic/electrostatic objective lens at the heart of the new instrument. The imaging aberrations of this new type of lens are said to decrease with decreasing beam energy. In addition, any beam crossover between the electron source (Schottky FE-gun) and the sample has been eliminated in order to minimize stochastic electron-electron interactions resulting in unwanted beam widening and

broadening of the beam energy spread. The secondary electron signal emitted by the sample is detected by an annular detector situated above the final lens. The instrument is computer-controlled and suitable for remote operation.

HazeBuster, produced by VayTek, is **software that operates as a digital confocal microscope** (*Reader Service No. 109*). It is said to be ideal for researchers who need to keep costs down, but must have the low-noise, highly detailed images that are normally obtained from expensive confocal equipment. The program uses proprietary deconvolution algorithms and a point-



**New software for confocal-quality images.**

spread function (measured or calculated) to remove the haze from microscope images. Even images captured with a scanning confocal can be enhanced with HazeBuster. The Power Mac edition also provides a suite of tools that aid in quantitative analysis and in image processing.

## Intracellular imaging

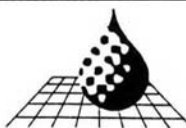
New **systems for intracellular ion measurement** are rolling off the production line at Applied Imaging (*Reader Service No. 110*). The QuantiCell 700, a spatial and temporal ration ion-imaging system, is a multiwavelength imaging system designed to make real-time imaging at up to 25 frames per second accessible to biomedical researchers operating on a budget. It is capable of imaging and calibrated measurement of ion-sensitive optical probes, such as the family of intracellular calcium probes Fura2, Indo1 and Fluo3,

and pH-sensitive fluorescent probes, including BCECF, SNARF and SNAFL, and membrane potential probes, such as Annexin and DioC5. It comes with a custom-designed frame grabber, a high-speed filter wheel controller and intensified CCD camera. The system's filter wheel offers not one but four wheels, with simultaneous operation on the excitation and emission ports. A neutral density wheel is designed to address the often difficult problem of balancing the light energy for ratiometric experiments, particularly those involving multiple probes. An array of image analysis tools is offered for obtaining fully calibrated, quantitative data from living cell experiments.

Version 3.1 of Kalcium-PC, Kinetic Imaging's **intracellular and digital video microscopy system**, is now shipping (*Reader Service No. 111*). Design changes include Astromed's 4100 high-speed, cooled CCD camera. The upgraded system also offers data rates to eight million pixels per second resolution to  $2,048 \times 2,048$  pixels and a 12-bit dynamic range. The company's enhanced multidimensional image manager provides a powerful route for handling multiwavelength, time-series images as single objects. Image disk technology is now available on single, high-performance AV disk drives. It provides a dedicated resource for acquisition, storage and management of massive, dynamic image data files. With Kalcium-PC, 2-, 4- and 9-Gbyte drives are available. The image disk technology also forms the basis of a new digital video microscopy system, Tempus-PC, which offers digital capture, editing and analysis of video at 25 or 30 f.p.s. □

*These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.*

## ADVERTISEMENTS



# Severn Biotech Ltd

Severn Biotech Ltd, Unit 2, Park Lane,  
Stourport Road, Kidderminster, Worcs DY11 6TJ, U.K.

**Custom DNA & RNA Oligo's**  
Probes, Primers, Antisense Oligo's, Ribonucleotides, 5' and 3' labelling with Biotin Fluorescein, Rhodamine, Modified bases purification etc.

**Complete Peptide Service**  
Small Scale Multiple Peptides 5-10mgs or larger scale 1000mgs using L and D Amino Acids, screening grade of HPLC purified. Phosphorylated, Acetylated Amidated and Protein Carrier Conjugated (KLH, PPD, etc.)

**Customer Services**  
Comprehensive Custom Service for DNA, RNA and Peptides from SBL

- DNA Synthesis
- RNA Synthesis
- Multi Peptide Synthesis (5-10mgs)
- Larger Scale Peptide Synthesis 1000mgs
- HPLC Purification
- Custom Conjugation to Carriers
- Peptide Sequencing / Mass Spec.

**For a Fast Efficient Service Reply Today**  
Telephone: (01562) 825286 Fax: (01562) 825284

Reader Service No. 80

## PACEMAKER®

Understanding the needs of the Life Scientist

Custom peptide, conjugate and antibody services

When your research depends on high quality at low cost turn to **PACEMAKER®**

All **PACEMAKER®** peptides, whether >70%, >80% or >95% purity, undergo hplc and mass spectrometric analysis.

Peptides are available in quantities from 1mg to 100mg and, with prices from **£115** per peptide and **£90** per conjugate, even the tightest budget is catered for.

Most importantly, you may come to us for advice. Based on many years experience and expertise we will try to find the best approach to fulfilling your custom peptide and antibody requirements.

**A F F I N I T Y**

Research Products Limited  
Monkhead Castle, Monkhead  
Evesham EX6 8HD, United Kingdom  
Tel: +44 (0) 1626 891010 Fax: +44 (0) 1626 891090

Reader Service No.70